GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 23:15:44 ; Search time 2325.7 Seconds
(without alignments)
3211.325 Million cell updates/sec

Title: US-09-303-518D-651
Perfect score: 4350
Sequence: 1 atgaaaacaaccgacaaacg.....aattaggctaccgctggtaa 4350

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4288	98.6	4350	20	AAZ12252	Neisseria meningit
	2	3555.8	81.7	4365	21	AAZ38917	Neisseria meningit
	3	3508	80.6	4374	20	AAZ12251	Neisseria meningit
	4	3508	80.6	4374	21	AAF21587	N. meningitidis pa
	5	3508	80.6	4374	21	AAA81302	N. meningitidis pa
	6	3508	80.6	26778	21	AAA81477	N. meningitidis pa
c	7	3508	80.6	349980	21	AAF21612	Neisseria meningit
c	8	3508	80.6	837096	21	AAA81489	N. meningitidis pa
	9	3498.4	80.4	4374	21	AAZ38918	Neisseria meningit
	10	3385.8	77.8	4407	20	AAZ12253	Neisseria gonorrho
	11	1348	31.0	4319	17	AAT17215	Adhesion and penet
	12	1048.2	24.1	2991	20	AAZ12250	Neisseria meningit
	13	790.8	18.2	891	21	AAA81401	N. meningitidis Me
	14	739.4	17.0	781	21	AAA81402	N. meningitidis Me
	15	690.8	15.9	1312	21	AAA81400	N. meningitidis Me
	16	206	4.7	969	18	AAT88142	H. influenzae Hap
c	17	191.4	4.4	30078	21	AAA81520	N. meningitidis pa
c	18	191.4	4.4	349980	21	AAF21608	Neisseria meningit
	19	177.4	4.1	4296	21	AAZ54326	Neisseria meningit
c	20	177.4	4.1	16526	21	AAA81472	N. meningitidis pa
	21	177.4	4.1	172325	21	AAF21613	Neisseria meningit
	22	177.4	4.1	349980	21	AAF21612	Neisseria meningit
	23	158	3.6	5091	11	AAQ06164	The iga gene, enco
	24	138.2	3.2	4113	22	ABA89172	Escherichia coli p
c	25	138.2	3.2	48254	22	ABA89141	Escherichia coli p
c	26	138.2	3.2	48345	22	ABA89142	Escherichia coli p
	27	122	2.8	4128	22	ABA88994	Escherichia coli p
c	28	122	2.8	7654	22	ABA88991	Escherichia coli p
	29	108.2	2.5	642	21	AAA81729	N. meningitidis pa
	30	52.2	1.2	10732	21	AAA10594	Gene encoding a su
	31	47.6	1.1	3489	21	AAA30290	Kaposi's sarcoma-a
	32	47.6	1.1	3489	22	AAF82901	Nucleotide sequenc
c	33	47.6	1.1	32207	20	AAV73805	KSHV LUR DNA (nucl
c	34	47.6	1.1	137507	19	AAV19941	KSHV long unique c
	35	43.8	1.0	4403765	22	AAI99683	Mycobacterium tube
	36	43.8	1.0	4411529	22	AAI99682	Mycobacterium tube
	37	43.6	1.0	2068	22	AAH98282	Rabbit EST-derived
	38	43	1.0	35133	21	AAA81518	N. meningitidis pa
	39	43	1.0	349980	21	AAF21611	Neisseria meningit
	40	43	1.0	837096	21	AAA81489	N. meningitidis pa
c	41	42.6	1.0	1089	21	AAA47543	Primate GPR2 seque
	42	42.2	1.0	400	19	AAV64545	M. tuberculosis im
	43	42.2	1.0	400	19	AAV44436	Mycobacterium tube
	44	42.2	1.0	400	20	AAZ19346	M. tuberculosis an
	45	42.2	1.0	400	20	AAZ19134	M. tuberculosis re
	46	41.6	1.0	6741	21	AAA10595	Gene encoding a su
c	47	41.4	1.0	4403765	22	AAI99683	Mycobacterium tube
c	48	41.2	0.9	272	21	AAC00277	Human secreted pro
c	49	41	0.9	4260	9	AAN81768	Sequence encoding
c	50	41	0.9	4411529	22	AAI99682	Mycobacterium tube
	51	40.6	0.9	4461	23	ABL14674	Drosophila melanog
	52	40.6	0.9	26879	23	AAS59524	Propionibacterium
	53	40.2	0.9	1461	23	ABL18553	Drosophila melanog
	54	40.2	0.9	4010	23	ABL18552	Drosophila melanog
	55	40	0.9	3741	21	AAZ39509	DNA sequence ID No
	56	40	0.9	58857	21	AAA58471	Nucleotide sequenc
	57	39.8	0.9	1068	21	AAA58928	cDNA encoding a Ne
	58	39.8	0.9	1098	24	AAS97227	Neisseria meningit
c	59	39.8	0.9	2716	23	ABL15073	Drosophila melanog
c	60	39.8	0.9	4380	9	AAN80222	Sequence of Mycoba
c	61	39.8	0.9	4380	19	AAV05708	Mycobacterium tube
c	62	39.6	0.9	567	21	AAA29550	HIV codon altered
	63	39.6	0.9	2885	23	ABL08459	Drosophila melanog
c	64	39.6	0.9	3848	23	ABL08136	Drosophila melanog
c	65	39.6	0.9	5004	23	ABL08458	Drosophila melanog
	66	39.4	0.9	2800	22	AAD07131	Canine retinitis p
	67	39.4	0.9	2803	22	AAD07132	Canine retinitis p
	68	39.4	0.9	2805	22	AAD07130	Canine retinitis p
	69	39.2	0.9	1439	19	AAV64533	M. tuberculosis im
	70	39.2	0.9	1439	19	AAV44424	Mycobacterium tube
	71	39.2	0.9	1439	20	AAZ19334	M. tuberculosis an
	72	39.2	0.9	1439	20	AAZ19122	M. tuberculosis re
	73	39.2	0.9	2408	20	AAV64112	Rhodobacter sphaer
	74	39.2	0.9	2408	20	AAV62125	R. sphaeroides Adh
	75	39.2	0.9	2408	22	AAF90615	Rhodobacter sphaer
	76	39.2	0.9	5975	15	AAQ55758	Klebsiella pneumon
	77	39.2	0.9	27541	22	AAD17185	Streptomyces nours
	78	39.2	0.9	37337	23	AAS59518	Propionibacterium
	79	39.2	0.9	125401	22	AAD17186	Streptomyces nours
	80	39	0.9	1072	23	AAS90738	DNA encoding novel
	81	39	0.9	1925	22	AAD18251	Sphingomonas elode
	82	39	0.9	2271	11	AAQ06844	Amylase gene from

	83	38.8	0.9	943	16	AAQ89539	Wheat gold toleran
	84	38.8	0.9	1628	16	AAQ80009	cDNA encoding cyst
	85	38.8	0.9	2294	18	AAT63246	Genomic sequence e
	86	38.8	0.9	2414	16	AAQ80010	Genomic DNA encodi
	87	38.8	0.9	3117	21	AAA49439	Cation channel gen
	88	38.8	0.9	4804	21	AAZ46086	cDNA encoding a re
	89	38.6	0.9	372	23	AAS75452	DNA encoding novel
	90	38.6	0.9	1596	23	ABL27423	Drosophila melanog
	91	38.6	0.9	1957	18	AAT93809	An artificial modi
c	92	38.6	0.9	3596	23	ABL27422	Drosophila melanog
c	93	38.6	0.9	4008	17	AAT14528	Rat adenylyl cycla
c	94	38.6	0.9	4008	21	AAA53919	Adenylyl cyclase t
c	95	38.6	0.9	5774	22	AAS45486	Chemically pretrea
	96	38.4	0.9	960	22	ABA01504	Codon modified hum
c	97	38.4	0.9	1404	21	AAZ54261	Neisseria meningit
	98	38.4	0.9	1621	19	AAV07163	Thermomonospora fu
	99	38.4	0.9	25509	21	AAA81467	N. meningitidis pa
	100	38.4	0.9	349980	21	AAF21607	Neisseria meningit

ALIGNMENTS

```
RESULT  1
AAZ12252
ID   AAZ12252 standard; DNA; 4350 BP.
XX
AC   AAZ12252;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria meningitidis strain A complete ORF1 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38824.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 364-365; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 4350 BP; 1225 A; 1161 C; 1060 G; 842 T; 62 other;

  Query Match             98.6%;  Score 4288;  DB 20;  Length 4350;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4350;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
Db       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60

Qy      61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
Db      61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120

Qy     121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
Db     121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180

Qy     181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
Db     181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240

Qy     241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
Db     241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300

Qy     301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
Db     301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360

Qy     361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
Db     361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420

Qy     421 gtgaaaagaaataattataagcctgacaattcacacccttacaacggcgattancatatg 480
Db     421 gtgaaaagaaataattataagcctgacaattcacacccttacaacggcgattancatatg 480

Qy     481 ccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatgagg 540
Db     481 ccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatgagg 540

Qy     541 gggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcaggacac 600
Db     541 gggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcaggacac 600

Qy     601 cactattggcgttatgatgatgacaaacacggcgatttatcctactccggcgcatggtta 660
Db     601 cactattggcgttatgatgatgacaaacacggcgatttatcctactccggcgcatggtta 660

Qy     661 attggcggcaatacacatatgcagggttggggaaataatggcgtanttagtttgagcggc 720
Db     661 attggcggcaatacacatatgcagggttggggaaataatggcgtanttagtttgagcggc 720

Qy     721 gatgtgcgccatgccaacgactatggccctatgccgattgcaggtgcggcaggcgacagc 780
Db     721 gatgtgcgccatgccaacgactatggccctatgccgattgcaggtgcggcaggcgacagc 780

Qy     781 ggttcgccaatgtttatttatgacaaaacaaacaataaatggctgctcaacggagtttta 840
Db     781 ggttcgccaatgtttatttatgacaaaacaaacaataaatggctgctcaacggagtttta 840

Qy     841 caaaccggctacccttattccggcagggaaaacggtttccagctgatacgcaaagattgg 900
Db     841 caaaccggctacccttattccggcagggaaaacggtttccagctgatacgcaaagattgg 900

Qy     901 ttctacgatgacatttacagaggcgatacacataccgtctnttttgaaccgcgcagtaac 960
Db     901 ttctacgatgacatttacagaggcgatacacataccgtctnttttgaaccgcgcagtaac 960

Qy     961 ggacatttttccttacatccaacaacaacggtacgggtacggtaacagaaaccaacgaa 1020
```

```
Db    961 ggacatttttcctttacatccaacaacaacggtacgggtacggtaacagaaaccaacgaa 1020
Qy   1021 aaggtntccaatccaaagcttaaagtacagacagtccgactgtttgacgaatctttgaat 1080
Db   1021 aaggtntccaatccaaagcttaaagtacagacagtccgactgtttgacgaatctttgaat 1080
Qy   1081 gaaactgataaagaaccagtttacgcggcaggggtgttaatcagtaccgtccaaggtta 1140
Db   1081 gaaactgataaagaaccagtttacgcggcaggggtgttaatcagtaccgtccaaggtta 1140
Qy   1141 aacaacggtgaaaacctttctttttatcgattacggcaacggcaaactcatcttatcaaac 1200
Db   1141 aacaacggtgaaaacctttctttttatcgattacggcaacggcaaactcatcttatcaaac 1200
Qy   1201 aacatcaaccaaggcgcgggcggtttgtattttgaaggtgattttacggtctcgcctgaa 1260
Db   1201 aacatcaaccaaggcgcgggcggtttgtattttgaaggtgattttacggtctcgcctgaa 1260
Qy   1261 aacaacgaaacgtggcaaggcgcgggcgttcatatcagtgaagacagtaccgttacttgg 1320
Db   1261 aacaacgaaacgtggcaaggcgcgggcgttcatatcagtgaagacagtaccgttacttgg 1320
Qy   1321 aaagtaaacggcgtggcaaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgtt 1380
Db   1321 aaagtaaacggcgtggcaaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgtt 1380
Qy   1381 caagccaaaggggaaaaccaaggctcgatcagcgtgggcgacggtacagtcattttggat 1440
Db   1381 caagccaaaggggaaaaccaaggctcgatcagcgtgggcgacggtacagtcattttggat 1440
Qy   1441 cagcaggcagacgataaaggcaaaaaacaagcctttagtgaaatcggcttgntcagcggc 1500
Db   1441 cagcaggcagacgataaaggcaaaaaacaagcctttagtgaaatcggcttgntcagcggc 1500
Qy   1501 aggggtacggtgcaactgaatgccgataatcagttcaaccccgacaaactctatttcggc 1560
Db   1501 aggggtacggtgcaactgaatgccgataatcagttcaaccccgacaaactctatttcggc 1560
Qy   1561 tttcgcggcggacgtttggatttaaacgggcattcgctttcgttccaccgtattcaaaat 1620
Db   1561 tttcgcggcggacgtttggatttaaacgggcattcgctttcgttccaccgtattcaaaat 1620
Qy   1621 accgatgaaggggcgatgattgncnatcataatgccacaacaacatccaccgttaccatt 1680
Db   1621 accgatgaaggggcgatgattgncnatcataatgccacaacaacatccaccgttaccatt 1680
Qy   1681 acagggaatgaaagtattacacaaccgagtggtaagaatatcaatagacttaattacagc 1740
Db   1681 acagggaatgaaagtattacacaaccgagtggtaagaatatcaatagacttaattacagc 1740
Qy   1741 aaagaaattgcctacaacggttggtttggcgagaaagatacgaccaaaacgaacgggcgg 1800
Db   1741 aaagaaattgcctacaacggttggtttggcgagaaagatacgaccaaaacgaacgggcgg 1800
Qy   1801 ctcaaccttgtttaccagcccgccgcagaagaccgcacccngctgctttccggcggaaca 1860
Db   1801 ctcaaccttgtttaccagcccgccgcagaagaccgcacccngctgctttccggcggaaca 1860
Qy   1861 aatttaaacggcaacatcacgcaaacaaacggcaaactgtttttcagcggcagaccgaca 1920
Db   1861 aatttaaacggcaacatcacgcaaacaaacggcaaactgtttttcagcggcagaccgaca 1920
Qy   1921 ccgcacgcctacaatcatttaggaagcgggtggtcaaaaatggaaggtatcccacaagga 1980
Db   1921 ccgcacgcctacaatcatttaggaagcgggtggtcaaaaatggaaggtatcccacaagga 1980
Qy   1981 gaaatcgtgtgggacaacgactggatcnaccgcacgtttaaagcggaaaatttccatatt 2040
Db   1981 gaaatcgtgtgggacaacgactggatcnaccgcacgtttaaagcggaaaatttccatatt 2040
Qy   2041 cagggcgggcaggcggtgatttcccgcaatgttgccaaagtggaaggcgattgncatttg 2100
Db   2041 cagggcgggcaggcggtgatttcccgcaatgttgccaaagtggaaggcgattgncatttg 2100
Qy   2101 agcaatcacgcccaagcagtttttggtgtcgcaccgcatcaaagccatacaatctgtaca 2160
Db   2101 agcaatcacgcccaagcagtttttggtgtcgcaccgcatcaaagccatacaatctgtaca 2160
Qy   2161 cgttcggactggacnggtctgacaaattgtgtcgaanaaancattaccgacgataaagtg 2220
Db   2161 cgttcggactggacnggtctgacaaattgtgtcgaanaaancattaccgacgataaagtg 2220
Qy   2221 attgcttcattgactaagacngacntnagcggcantgtnagnctnnccnatnacgntnnt 2280
Db   2221 attgcttcattgactaagacngacntnagcggcantgtnagnctnnccnatnacgntnnt 2280
Qy   2281 tnaaanctcncngggcntgcnncactnaanggcaatcttagtgcaaatggcgatacacgt 2340
Db   2281 tnaaanctcncngggcntgcnncactnaanggcaatcttagtgcaaatggcgatacacgt 2340
Qy   2341 tatacagtcagccacaacgccacccaaaacggcaaccttagcctcgtgggcaatgcccaa 2400
Db   2341 tatacagtcagccacaacgccacccaaaacggcaaccttagcctcgtgggcaatgcccaa 2400
Qy   2401 gcaacatttaatcaagccacattaaacggcaacncatcggnttcgggcaatgcttcattt 2460
Db   2401 gcaacatttaatcaagccacattaaacggcaacncatcggnttcgggcaatgcttcattt 2460
Qy   2461 aatctaagcaacaacgccgcacaaaacggcagtctgacgctttccgacaacgctaaggca 2520
Db   2461 aatctaagcaacaacgccgcacaaaacggcagtctgacgctttccgacaacgctaaggca 2520
Qy   2521 aacgtaagccattccgcactcaacggcaatgtctccctagccgataaggcagtattccat 2580
Db   2521 aacgtaagccattccgcactcaacggcaatgtctccctagccgataaggcagtattccat 2580
Qy   2581 tttgaaaacagccgctttaccggacaactcagcggcagcaagganacagcattacactta 2640
Db   2581 tttgaaaacagccgctttaccggacaactcagcggcagcaagganacagcattacactta 2640
Qy   2641 aaagacagcgaatggacgctgccgtcaggcacggaattaggcaatttaaaccttgacaac 2700
Db   2641 aaagacagcgaatggacgctgccgtcaggcacggaattaggcaatttaaaccttgacaac 2700
Qy   2701 gccaccattacactcaattccgcctatcgccacgatgctgcaggcgcgcaaaccggcagn 2760
Db   2701 gccaccattacactcaattccgcctatcgccacgatgctgcaggcgcgcaaaccggcagn 2760
Qy   2761 gtgtcagacacgccgcgccgccgttcgcgccgttccctattatccgttacaccgccaact 2820
Db   2761 gtgtcagacacgccgcgccgccgttcgcgccgttccctattatccgttacaccgccaact 2820
Qy   2821 tcggtagaatcccgtttcaacacgctgacggtaaacggcaaattgaacngtcaaggaaca 2880
Db   2821 tcggtagaatcccgtttcaacacgctgacggtaaacggcaaattgaacngtcaaggaaca 2880
Qy   2881 ttccgctttatgtcggaactcttcggctaccgaagcgacaaattgaagctggcggaaagt 2940
Db   2881 ttccgctttatgtcggaactcttcggctaccgaagcgacaaattgaagctggcggaaagt 2940
Qy   2941 tccgaaggnacttacaccttggcggtcaacaataccggcaacgaacccgtaagcctcgat 3000
Db   2941 tccgaaggnacttacaccttggcggtcaacaataccggcaacgaacccgtaagcctcgat 3000
Qy   3001 caattgacggtagtggaagggaaagacaacaaaccgctgtccgaaaaccttaatttcacc 3060
Db   3001 caattgacggtagtggaagggaaagacaacaaaccgctgtccgaaaaccttaatttcacc 3060
Qy   3061 ctgcaaaacgaacacgtcgatgccggcgcgtggcgttaccaactcatccgcaaagacggc 3120
Db   3061 ctgcaaaacgaacacgtcgatgccggcgcgtggcgttaccaactcatccgcaaagacggc 3120
Qy   3121 gagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactcggcaaggca 3180
Db   3121 gagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactcggcaaggca 3180
```

```
Qy   3181 gaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctgattgcggcc 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 gaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctgattgcggcc 3240

Qy   3241 gggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcnggcaggcggg 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 gggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcnggcaggcggg 3300

Qy   3301 gaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcggataaagac 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 gaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcggataaagac 3360

Qy   3361 agcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgccttcccccgc 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 agcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgccttcccccgc 3420

Qy   3421 gcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacctcaaccccaa 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 gcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacctcaaccccaa 3480

Qy   3481 ccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattttccgccacg 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 ccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattttccgccacg 3540

Qy   3541 ctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaagaccgccgc 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaagaccgccgc 3600

Qy   3601 aacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaagatttccgc 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 aacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaagatttccgc 3660

Qy   3661 gcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctcggcagcggg 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 gcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctcggcagcggg 3720

Qy   3721 cgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgacggcatcggc 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 cgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgacggcatcggc 3780

Qy   3781 aactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggcaggttcgac 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 aactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggcaggttcgac 3840

Qy   3841 atcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggcatcggaggc 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 atcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggcatcggaggc 3900

Qy   3901 aaaatccgccgccgcgtgctgcattacggcattcaggcacgataccgcgccggtttcggc 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 aaaatccgccgccgcgtgctgcattacggcattcaggcacgataccgcgccggtttcggc 3960

Qy   3961 ggattcggcatcgaaccgtacatcggcgcaacgcgctatttcgtccaaaaagcggattac 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 ggattcggcatcgaaccgtacatcggcgcaacgcgctatttcgtccaaaaagcggattac 4020

Qy   4021 cgctacgaaaacgtcaatatcgccacccccggtcttgcgttcaaccgntaccgngcgggc 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 cgctacgaaaacgtcaatatcgccacccccggtcttgcgttcaaccgntaccgngcgggc 4080

Qy   4081 attaaggcagattattcattcaaaccggcgcaacacatntccatcacnccttatttnagc 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 attaaggcagattattcattcaaaccggcgcaacacatntccatcacnccttatttnagc 4140

Qy   4141 ctgtcctataccgatgccgcttcgggcaaagtccgaacacgcgtcaataccgcngtattg 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ctgtcctataccgatgccgcttcgggcaaagtccgaacacgcgtcaataccgcngtattg 4200

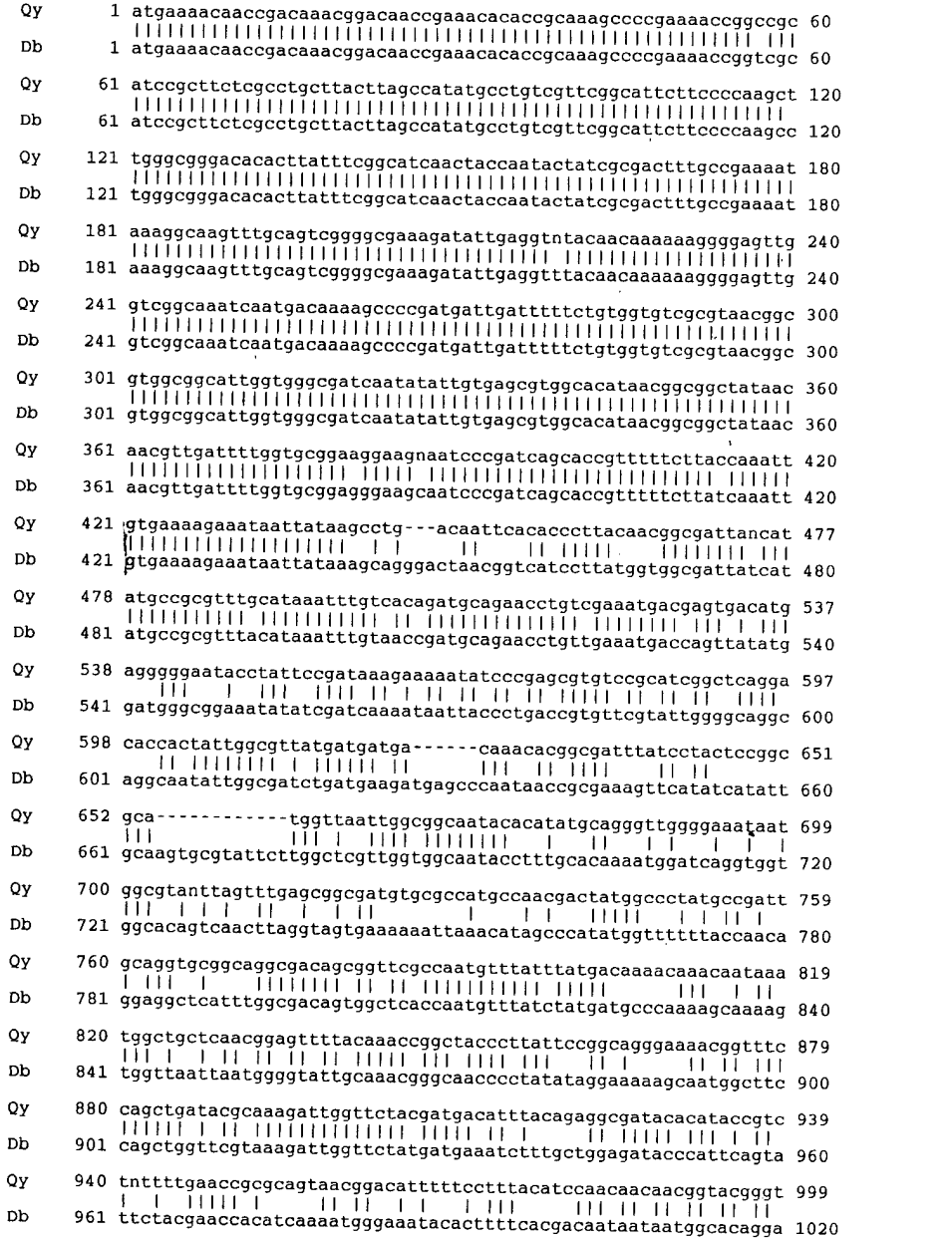
Qy   4201 gctcaggatttcggcaaaacccgcagtgcggaatggggcgtaaacgccgaaatcaaaggt 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 gctcaggatttcggcaaaacccgcagtgcggaatggggcgtaaacgccgaaatcaaaggt 4260

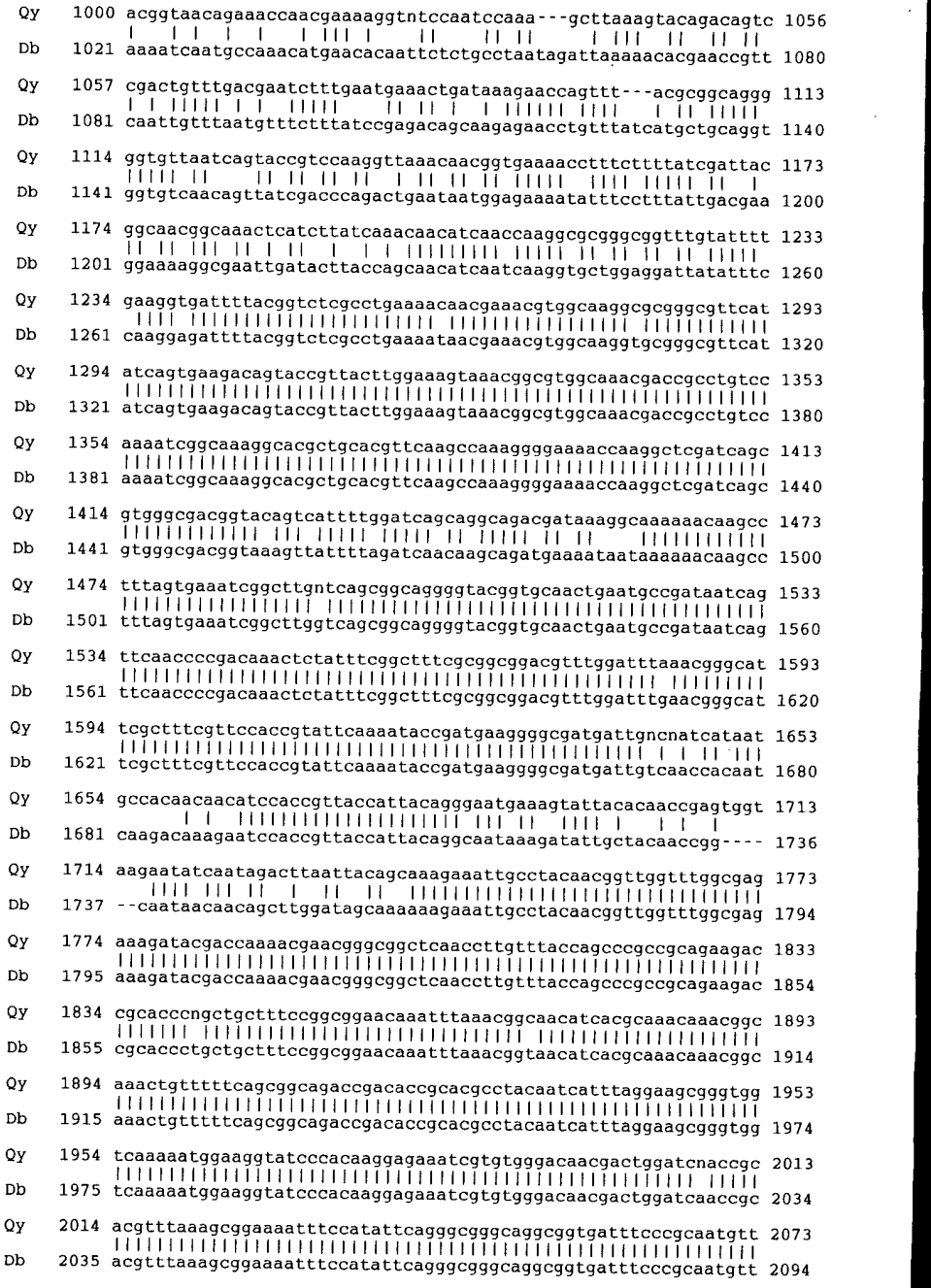
Qy   4261 ttcacgctgtccntccacgctgccgccgccaaaggnccgcaactggaagcgcaacacagc 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 ttcacgctgtccntccacgctgccgccgccaaaggnccgcaactggaagcgcaacacagc 4320

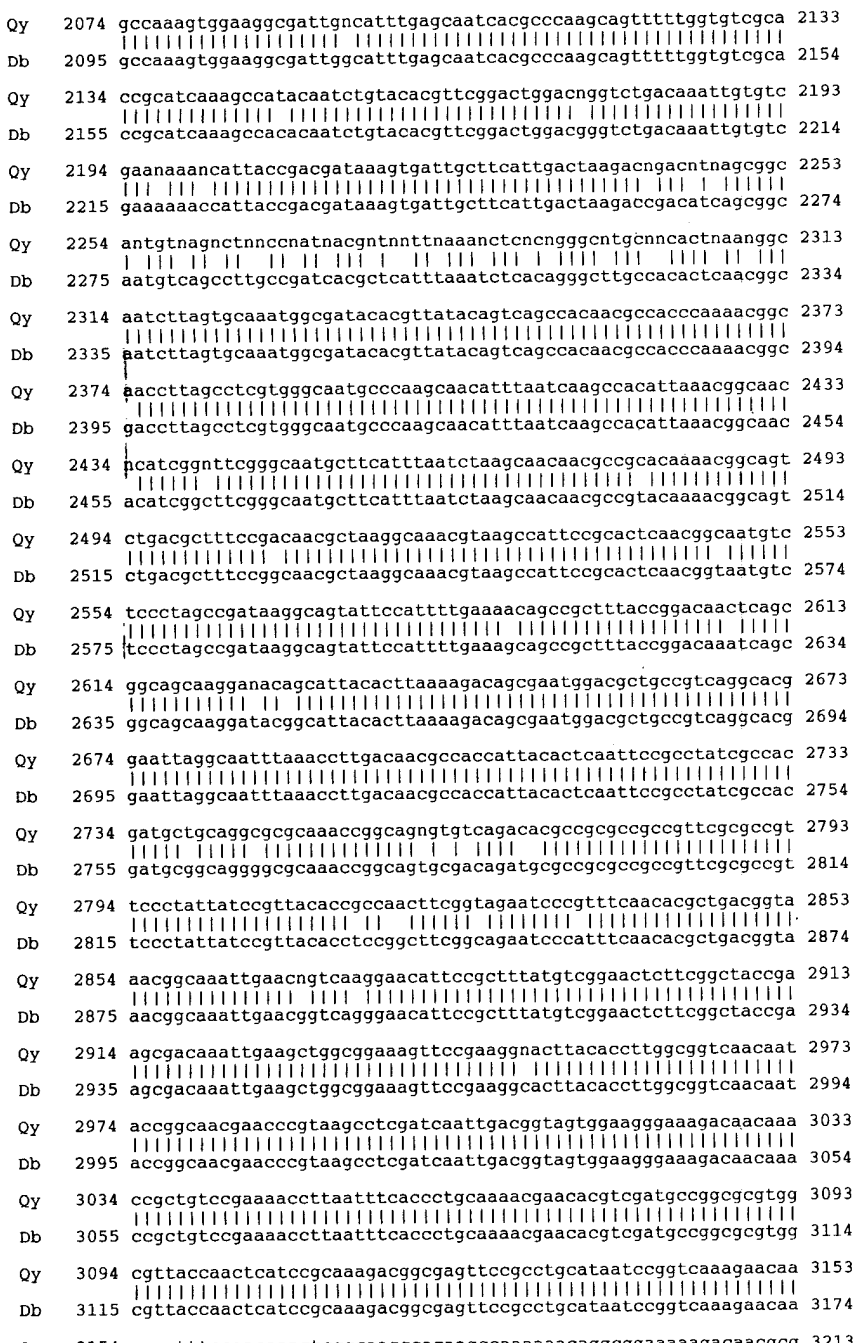
Qy   4321 gcgggcatcaaattaggctaccgctggtaa 4350
          |||||||||||||||||||||||||||||
Db   4321 gcgggcatcaaattaggctaccgctggtaa 4350
```

RESULT 2
AAZ38917
ID AAZ38917 standard; DNA; 4365 BP.
XX
AC AAZ38917;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain ATCC13090 BASB006 nucleotide sequence.
XX
KW Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KW vaccine; antibiotic; upper respiratory tract infection; meningitis;
KW invasive bacterial disease; bacteraemia; screening; antibacterial; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1..4365
FT /*tag= a
FT /product= "BASB006"
XX
PN WO9955873-A2.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP02766.
XX
PR 24-APR-1998; 98GB-0008866.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Thonnard J;
XX
DR WPI; 2000-052810/04.
DR P-PSDB; AAY56621.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitis used
PT to prepare vaccines against bacterial infections -
XX
PS Claim 11; Page 84-86; 103pp; English.
XX
CC The present sequence encodes BASB006 isolated from Neisseria meningitidis
CC strain ATCC13090. BASB006 polynucleotides and polypeptides may be
CC employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of disease, staging of disease, or
CC determining response of an infectious organism to drugs. The
CC polynucleotides may be used as a source for hybridisation probes, and for
CC screening of genetic mutations, serotype, organism or strain
CC identification, identification of mutation in BASB013 sequences, and as
CC components of arrays which are useful for diagnostic and prognostic
CC purposes. The polypeptides can be used to produce antibodies. The
CC polypeptides can also be used in vaccine formulations, and to identify
CC agonists and antagonists. The polypeptides, antibodies, agonists and
CC antagonists (which are bacteristatic) are used for the treatment and
CC prevention of diseases such as upper respiratory tract infection,
CC invasive bacterial diseases such as bacteraemia and meningitis, and for
CC the development and screening of antibacterial drugs. They are also used
CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC prevent tissue damage and/or block the normal progression of pathogenesis
CC in infections initiated other than by the implantation of in-dwelling
CC devices or by other surgical techniques.
XX
SQ Sequence 4365 BP; 1267 A; 1149 C; 1066 G; 883 T; 0 other;

Query Match 81.7%; Score 3555.8; DB 21; Length 4365;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 3894; Conservative 0; Mismatches 444; Indels 39; Gaps 7;






Qy 2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagtttttggtgtcgca 2133
Db 2095 gccaaagtggaaggcgattggcatttgagcaatcacgcccaagcagtttttggtgtcgca 2154
Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193
Db 2155 ccgcatcaaagccacacaatctgtacacgttcggactggacgggtctgacaaattgtgtc 2214
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 2215 gaaaaaaccattaccgacgataaagtgattgcttcattgactaagaccgacatcagcggc 2274
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 2275 aatgtcagccttgccgatcacgctcatttaaatctcacagggcttgccacactcaacggc 2334
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 2335 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2394
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db 2395 gaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2454
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 2455 acatcggcttcgggcaatgcttcatttaatctaagcaacaacgccgtacaaaacggcagt 2514
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 2515 ctgacgctttccggcaacgctaaggcaaacgtaagccattccgcactcaacggtaatgtc 2574
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 2575 tccctagccgataaggcagtattccattttgaaagcagccgctttaccggacaaatcagc 2634
Qy 2614 ggcagcaaggcanacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 2635 ggcagcaaggatacggcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2694
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 2695 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2754
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcgcgccgt 2793
Db 2755 gatgcggcaggggcgcaaaccggcagtgcgacagatgcgccgcgccgccgttcgcgccgt 2814
Qy 2794 tccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggta 2853
Db 2815 tccctattatccgttacacctccggcttcggcagaatcccatttcaacacgctgacggta 2874
Qy 2854 aacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttcggctaccga 2913
Db 2875 aacggcaaattgaacggtcagggaacattccgctttatgtcggaactcttcggctaccga 2934
Qy 2914 agcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcggtcaacaat 2973
Db 2935 agcgacaaattgaagctggcggaaagttccgaaggcacttacaccttggcggtcaacaat 2994
Qy 2974 accggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaagacaacaaa 3033
Db 2995 accggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaagacaacaaa 3054
Qy 3034 ccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgccggcgcgtgg 3093
Db 3055 ccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgccggcgcgtgg 3114
Qy 3094 cgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtcaaagaacaa 3153
Db 3115 cgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtcaaagaacaa 3174
Qy 3154 gagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcg 3213
Db 3175 gagctttccgacaaactcggcaaggcagaagccaaaaaacaggcgggaaaagacaacgcg 3234
Qy 3214 caaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagacagaaagcgtt 3273
Db 3235 caaagccttgacgcgctgattgcggccgggcgcgatgccgtcgaaaagacagaaagcgtt 3294
Qy 3274 gccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcggaggaagag 3333
Db 3295 gccgaaccggcccggcaggcaggcggggaaaatgtcggcattatgcaggcggaggaagag 3354
Qy 3334 aaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaagcggaaacc 3393
Db 3355 aaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaagggaaaacc 3414
Qy 3394 cggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttgccgcaaccg 3453
Db 3415 cggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttgccgcaaccg 3474
Qy 3454 cagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgttatgccaat 3513
Db 3475 cagccccaaccg------caaccccaaccgcagcgcgacctgatcagccgttatgccaat 3528
Qy 3514 agcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacaggacgaattg 3573
Db 3529 agcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacaggacgaatta 3588
Qy 3574 gaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatccggnacacc 3633
Db 3589 gaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatccgggacacc 3648
Qy 3634 aaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaaatc 3693
Db 3649 aaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaaatc 3708
Qy 3694 ggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggacc 3753
Db 3709 ggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggacc 3768
Qy 3754 gaaaacanctcgacgacggcatcggcaactcggcacggcttgcccacggcgccgttttc 3813
Db 3769 gaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgttttc 3828
Qy 3814 gggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcagc 3873
Db 3829 gggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcagc 3888
Qy 3874 ggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcatt 3933
Db 3889 ggcagtctttcagacgacatcggaagcaaaatccgccgccgcgtgctgcattacggcatt 3948
Qy 3934 caggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaacg 3993
Db 3949 caggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatcggcgcaacg 4008
Qy 3994 cgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccggt 4053
Db 4009 cgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccggc 4068
Qy 4054 cttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaaccggcgcaa 4113
Db 4069 cttgcgttcaaccgctaccgcgcgggcattaaggcagattattcattcaaaccggcgcaa 4128
Qy 4114 cacatntccatcacnccttatttnagcctgtcctataccgatgccgcttcgggcaaagtc 4173
Db 4129 cacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcgggcaaagtc 4188
Qy 4174 cgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgcagtgcggaa 4233
Db 4189 cgaacgcgcgtcaataccgccgtattggctcaggatttcggcaaaacccgcagtgcggaa 4248
Qy 4234 tggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgccgccgccaaa 4293

```
Db   4249 tggggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgccgccgccaaa 4308

Qy   4294 ggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggtaa 4350
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 ggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggtaa 4365


RESULT   3
AAZ12251
ID   AAZ12251 standard; DNA; 4374 BP.
XX
AC   AAZ12251;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Neisseria meningitidis complete ORF1 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB01665.
XX
PR   01-SEP-1998;   98GB-0019016.
PR   06-NOV-1997;   97GB-0023516.
PR   14-NOV-1997;   97GB-0024190.
PR   18-NOV-1997;   97GB-0024386.
PR   27-NOV-1997;   97GB-0025158.
PR   10-DEC-1997;   97GB-0026147.
PR   14-JAN-1998;   98GB-0000759.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38823.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection
XX
PS   Claim 9; Page 360-361; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC   (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC   antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC   fragments, their nucleic acids and antibodies are used for diagnosis,
CC   prevention (as vaccines) or treatment of Neisseria infections,
CC   such as meningitis, septicaemia and gonorrhea. Both organisms
CC   are closely related. Fragments of the nucleic acids are useful
CC   as hybridisation probes and antisense reagents.
XX
SQ   Sequence 4374 BP; 1266 A; 1163 C; 1066 G; 879 T; 0 other;


  Query Match             80.6%;  Score 3508;  DB 20;  Length 4374;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches 3876;  Conservative    0;  Mismatches  462;  Indels   48;  Gaps    8;

Qy      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60

Qy     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagcc 120

Qy    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180

Qy    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 240

Qy    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300

Qy    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360

Qy    361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
          |||||||||||||||||||||||| |||||||||| || |||||| |||| |||||
Db    361 aacgttgattttggtgcggaaggaagaaatcccgatcaacatcgttttacttataaaatt 420

Qy    421 gtgaaaagaaataattataagcctgacaattca---caccttacaacggcgattancat 477
          |||||| | |||||||||||  | |  | |  |   || |||||  ||||||||| |||
Db    421 gtgaaacggaataattataaagcagggactaaaggccatccttatggcggcgattatcat 480

Qy    478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
          |||||||||||||||||||||||||||||||||||||||| |||||||| ||| | |||
Db    481 atgccgcgtttgcataaatttgtcacagatgcagaacctgttgaaatgaccagttatatg 540

Qy    538 aggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
            |||   |  |||  |||| || | || || || || |||| || || ||  ||||
Db    541 gatgggcggaaatatatcgatcaaaataattaccctgaccgtgttcgtattggggcaggc 600

Qy    598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
             || |||||||| | |||||| ||      |||  || ||||    || ||
Db    601 aggcaatattggcgatctgatgaagatgagcccaataaccgcgaaagttcatatcatatt 660

Qy    652 gca------------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
          |||             ||| |  |||| ||||||||   |   ||   | |  |  |  |
Db    661 gcaagtgcgtattcttggctcgttggtggcaatacctttgcacaaaatggatcaggtggt 720

Qy    700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
          |||  | | |  ||  |  | ||         |    | |   |||||   | | || |
Db    721 ggcacagtcaacttaggtagtgaaaaaattaaacatagcccatatggttttttaccaaca 780

Qy    760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
          | |||  |    |||||||| || || ||||||||||| |||||      |||  | ||
Db    781 ggaggctcatttggcgacagtggctcaccaatgtttatctatgatgcccaaaagcaaaag 840

Qy    820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
          ||| |  | || || || || ||||| ||| |||| |||   || |     || || |||
Db    841 tggttaattaatggggtattgcaaacgggcaacccctatataggaaaaagcaatggcttc 900

Qy    880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
          |||||| | || |||||||||||||||| ||||| || |     || ||||| ||| | ||
Db    901 cagctggttcgtaaagattggttctatgatgaaatctttgctggagatacccattcagta 960

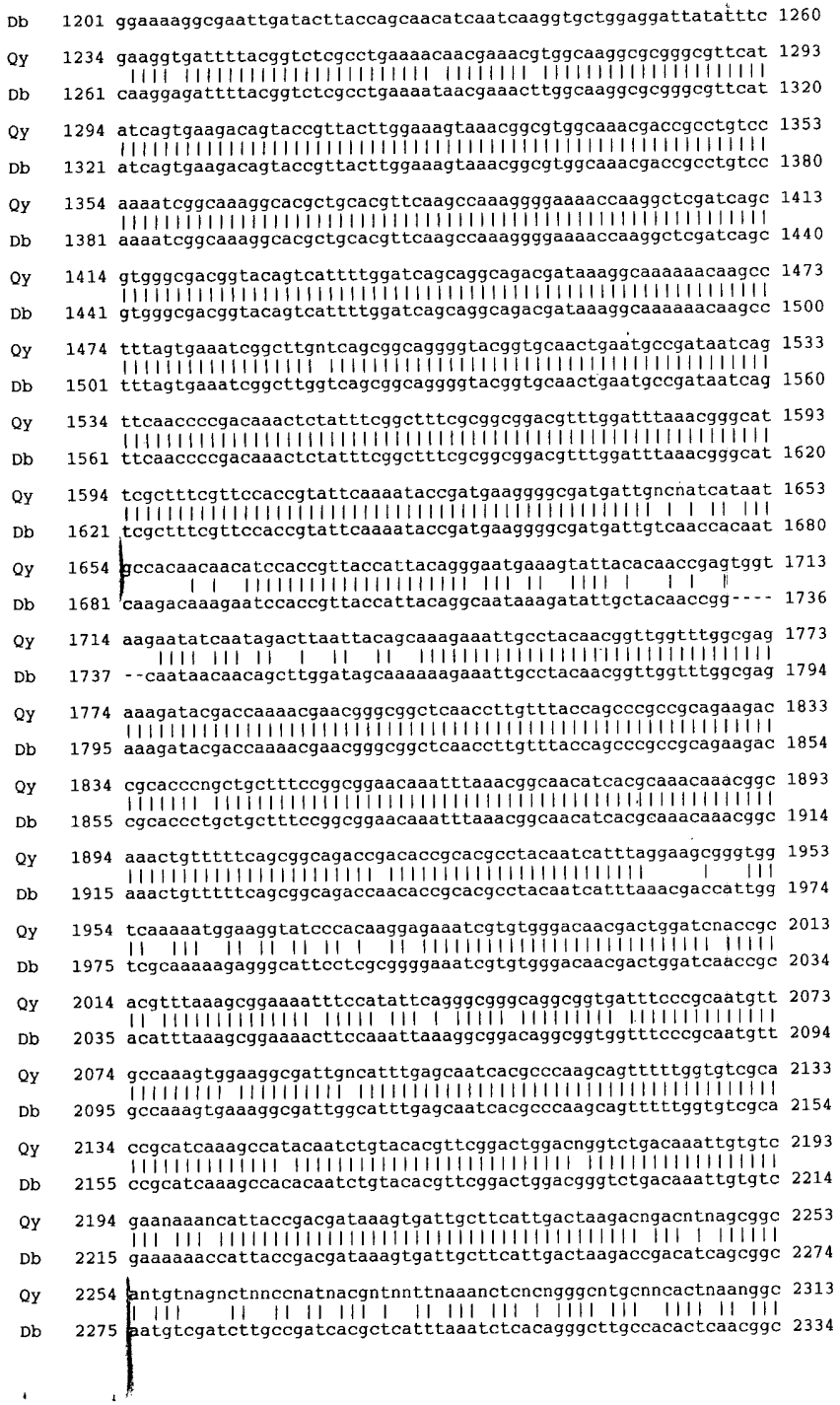
Qy    940 tnttttgaaccgcgcagtaacggacatttttccttttacatccaacaacaacggtacgggt 999
          |  |  ||||| ||    || ||  | |  || |||     | | || || || || ||
Db    961 ttctacgaaccacgtcaaaatgggaaatactcttttaacgacgataataatggcacagga 1020

Qy   1000 acggtaacagaaaccaacgaaaaggtntccaatccaaa---gcttaaagtacagacagtc 1056
          |   | |  |  |  | ||| |    ||    || ||     | |||  ||  || ||
Db   1021 aaaatcaatgccaaacatgaacacaattctctgcctaatagattaaaaacacgaaccgtt 1080

Qy   1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
          | | ||||| | |  |||||     || || |  | |||||| ||||   | || |||||
Db   1081 caattgtttaatgtttctttatccgagacagcaagagaacctgtttatcatgctgcaggt 1140

Qy   1114 ggtgttaatcagtaccgtccaaggttaaacaacggtgaaaaccttttcttttatcgattac 1173
          |||||  ||    || || || ||  | || || || |||||  |||| |||||  ||  |
Db   1141 ggtgtcaacagttatcgacccagactgaataatggagaaaatatttccttttattgacgaa 1200

Qy   1174 ggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttt 1233
          || || ||| || | ||  |  | | |||||||||| ||||| || || || || |||||
```


Db 1201 ggaaaaggcgaattgatacttaccagcaacatcaatcaaggtgctggaggattatatttc 1260
Qy 1234 gaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcat 1293
Db 1261 caaggagattttacggtctcgcctgaaaataacgaaacttggcaaggcgcgggcgttcat 1320
Qy 1294 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1353
Db 1321 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1380
Qy 1354 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1413
Db 1381 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1440
Qy 1414 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1473
Db 1441 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1500
Qy 1474 tttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcag 1533
Db 1501 tttagtgaaatcggcttggtcagcggcaggggtacggtgcaactgaatgccgataatcag 1560
Qy 1534 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1593
Db 1561 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1620
Qy 1594 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcataat 1653
Db 1621 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgtcaaccacaat 1680
Qy 1654 gccacaacaacatccaccgttaccattacagggaatgaaagtattacacaaccgagtggt 1713
Db 1681 caagacaaagaatccaccgttaccattacaggcaataaagatattgctacaaccgg---- 1736
Qy 1714 aagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtttggcgag 1773
Db 1737 --caataacaacagcttggatagcaaaaaagaaattgcctacaacggttggtttggcgag 1794
Qy 1774 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1833
Db 1795 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1854
Qy 1834 cgcacccngctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1893
Db 1855 cgcaccctgctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1914
Qy 1894 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1953
Db 1915 aaactgtttttcagcggcagaccaacaccgcacgcctacaatcatttaaacgaccattgg 1974
Qy 1954 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgactggatcnaccgc 2013
Db 1975 tcgcaaaaagagggcattcctcgcggggaaatcgtgtgggacaacgactggatcaaccgc 2034
Qy 2014 acgtttaaagcggaaaatttccatattcagggcgggcaggcggtgatttcccgcaatgtt 2073
Db 2035 acatttaaagcggaaaacttccaaattaaaggcggacaggcggtggtttcccgcaatgtt 2094
Qy 2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagtttttggtgtcgca 2133
Db 2095 gccaaagtgaaaggcgattggcatttgagcaatcacgcccaagcagtttttggtgtcgca 2154
Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193
Db 2155 ccgcatcaaagccacacaatctgtacacgttcggactggacgggtctgacaaattgtgtc 2214
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 2215 gaaaaaaccattaccgacgataaagtgattgcttcattgactaagaccgacatcagcggc 2274
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 2275 aatgtcgatcttgccgatcacgctcatttaaatctcacagggcttgccacactcaacggc 2334
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 2335 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2394
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db 2395 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2454
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 2455 acatcggcttcgggcaatgcttcatttaatctaagcgaccacgccgtacaaaacggcagt 2514
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 2515 ctgacgctttccggcaacgctaaggcaaacgtaagccattccgcactcaacggtaatgtc 2574
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 2575 tccctagccgataaggcagtattccattttgaaagcagccgctttaccggacaaatcagc 2634
Qy 2614 ggcagcaaggganacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 2635 ggcggcaaggatacggcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2694
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 2695 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2754
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgccgttcg------ 2787
Db 2755 gatgcggcaggggcgcaaaccggcagtgcgacagatgcgccgcgccgccgttcgcgccgt 2814
Qy 2788 ---cgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2844
Db 2815 tcgcgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2874
Qy 2845 ctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttc 2904
Db 2875 ctgacggtaaacggcaaattgaacggtcagggaacattccgctttatgtcggaactcttc 2934
Qy 2905 ggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcg 2964
Db 2935 ggctaccgcagcgacaaattgaagctggcggaaagttccgaaggcacttacaccttggcg 2994
Qy 2965 gtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaa 3024
Db 2995 gtcaacaataccggcaacgaacctgcaagcctcgaacaattgacggtagtggaagggaaaa 3054
Qy 3025 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3084
Db 3055 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3114
Qy 3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
Db 3115 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3174
Qy 3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204
Db 3175 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3234
Qy 3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
Db 3235 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgtcgaaaagaca 3294
Qy 3265 gaaagcgttgccgaaccggcccggcnggcaggcgggggaaaatgtcggcattatgcaggcg 3324
Db 3295 gaaagcgttgccgaaccggcccggcaggcaggcgggggaaaatgtcggcattatgcaggcg 3354
Qy 3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
Db 3355 gaggaagagaaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaa 3414

```
Qy   3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
          ||||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||
Db   3415 gcggaaacccggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttg 3474

Qy   3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
          ||||||| ||| ||||||||      || ||||||||||||||||||||| ||||||
Db   3475 ccgcaactgcaaccccaaccg------cagccccaaccgcagcgcgacctgatcagccgt 3528

Qy   3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3529 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3588

Qy   3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
          |||||||| |||||||| ||||||||||||||||||||| |||||||||||| |||||
Db   3589 gacgaattagaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatc 3648

Qy   3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3649 cgggacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3708

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3709 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3768

Qy   3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3769 aaccggaccgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggc 3828

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
          |||||||||||||||||||||| ||||||| |||||||||||||| ||||||||||||
Db   3829 gccgttttcgggcaatacggcatcgacaggttctacatcggcatcagcgcgggcgcgggt 3888

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
          ||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||||
Db   3889 tttagcagcggcagcctttcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3948

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   3949 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatc 4008

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4068

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
          |||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db   4069 acccccggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaa 4128

Qy   4105 ccggcgcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcg 4164
          |||||||||||||| |||||||| |||||||| |||||||||||||||||||||||||||
Db   4129 ccggcgcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcg 4188

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   4189 ggcaaagtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgc 4248

Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   4249 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtcctccacgctgcc 4308

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 gccgccaaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4368

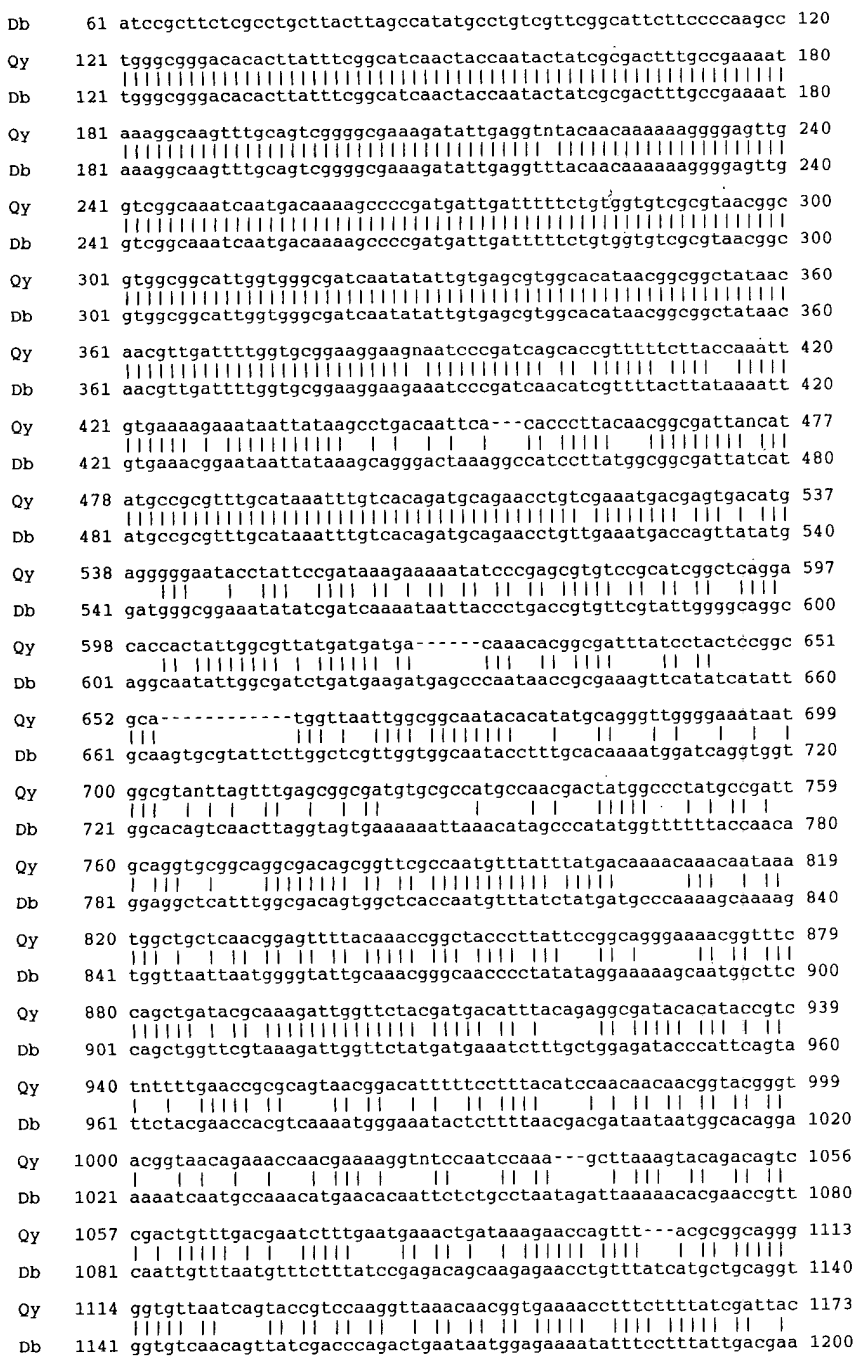
Qy   4345 tggtaa 4350
          ||||||
Db   4369 tggtaa 4374
```

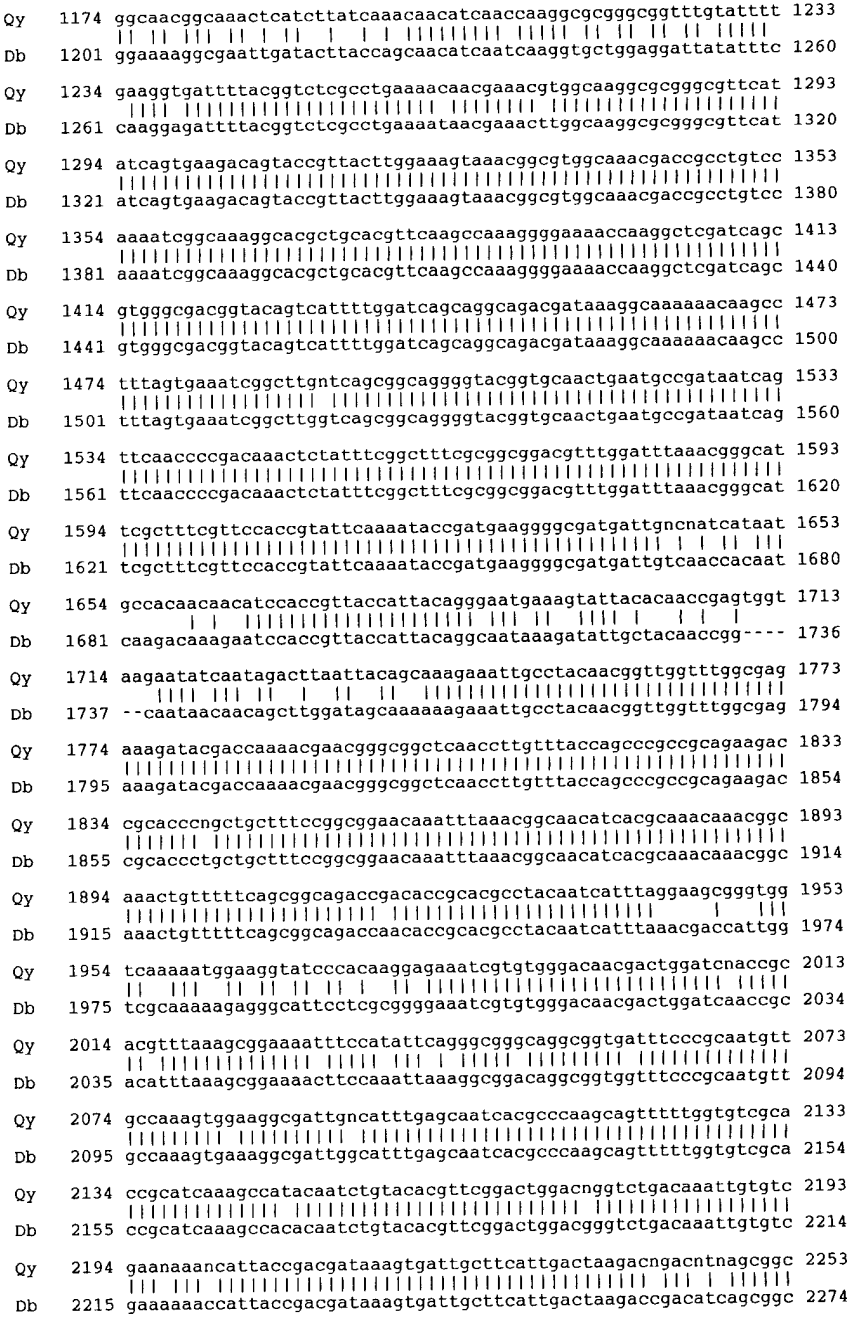
```
RESULT   4
AAF21587
ID   AAF21587 standard; DNA; 4374 BP.
XX
AC   AAF21587;
XX
DT   13-MAR-2001  (first entry)
XX
DE   N. meningitidis partial DNA sequence orf1-1.seq SEQ ID NO:86.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
DR   P-PSDB; AAB58592.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Example 1; Page 112-114; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 4374 BP; 1266 A; 1163 C; 1066 G; 879 T; 0 other;
```

```
  Query Match             80.6%;  Score 3508;  DB 21;  Length 4374;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches 3876;  Conservative    0;  Mismatches  462;  Indels   48;  Gaps    8;

Qy       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60

Qy      61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```





```
Qy   2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db   2275 aatgtcgatcttgccgatcacgctcatttaaatctcacagggcttgccacactcaacggc 2334

Qy   2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db   2335 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2394

Qy   2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db   2395 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2454

Qy   2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db   2455 acatcggcttcgggcaatgcttcatttaatctaagcgaccacgccgtacaaaacggcagt 2514

Qy   2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db   2515 ctgacgctttccggcaacgctaaggcaaacgtaagccattccgcactcaacggtaatgtc 2574

Qy   2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db   2575 tccctagccgataaggcagtattccattttgaaagcagccgctttaccggacaaatcagc 2634

Qy   2614 ggcagcaaggganacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db   2635 ggcggcaaggatacggcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2694

Qy   2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db   2695 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2754

Qy   2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcg------ 2787
Db   2755 gatgcggcaggggcgcaaaccggcagtgcgacagatgcgccgcgccgccgttcgccgccgt 2814

Qy   2788 ---cgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2844
Db   2815 tcgcgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2874

Qy   2845 ctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttc 2904
Db   2875 ctgacggtaaacggcaaattgaacggtcagggaacattccgctttatgtcggaactcttc 2934

Qy   2905 ggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcg 2964
Db   2935 ggctaccgcagcgacaaattgaagctggcggaaagttccgaaggcacttacaccttggcg 2994

Qy   2965 gtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaa 3024
Db   2995 gtcaacaataccggcaacgaacctgcaagcctcgaacaattgacggtagtggaaggaaaa 3054

Qy   3025 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3084
Db   3055 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3114

Qy   3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
Db   3115 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3174

Qy   3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204
Db   3175 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3234

Qy   3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
Db   3235 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgtcgaaaagaca 3294

Qy   3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
Db   3295 gaaagcgttgccgaaccggcccggcaggcaggcggggaaaatgtcggcattatgcaggcg 3354

Qy   3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
Db   3355 gaggaagagaaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaa 3414

Qy   3385 gcggaaaccccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
Db   3415 gcggaaaccccggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttg 3474

Qy   3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
Db   3475 ccgcaactgcaaccccaaccg------cagccccaaccgcagcgcgacctgatcagccgt 3528

Qy   3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
Db   3529 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3588

Qy   3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
Db   3589 gacgaattagaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatc 3648

Qy   3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
Db   3649 cgggacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3708

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
Db   3709 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3768

Qy   3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
Db   3769 aaccggaccgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggc 3828

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
Db   3829 gccgttttcgggcaatacggcatcgacaggttctacatcggcatcagcgcgggcgcgggt 3888

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
Db   3889 tttagcagcggcagcctttcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3948

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
Db   3949 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatc 4008

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
Db   4009 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4068

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
Db   4069 acccccggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaa 4128

Qy   4105 ccggcgcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcg 4164
Db   4129 ccggcgcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcg 4188

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
Db   4189 ggcaaagtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgc 4248

Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
Db   4249 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgcc 4308

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
Db   4309 gccgccaaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4368

Qy   4345 tggtaa 4350
Db   4369 tggtaa 4374

RESULT 5
```

```
AAA81302
ID   AAA81302 standard; DNA; 4374 BP.
XX
AC   AAA81302;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence orf1-1.seq SEQ ID NO:1046.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
DR   P-PSDB; AAB25662.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Example 1; Page 112-113; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 4374 BP; 1266 A; 1163 C; 1066 G; 879 T; 0 other;

  Query Match             80.6%;  Score 3508;  DB 21;  Length 4374;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches 3876;  Conservative    0;  Mismatches  462;  Indels   48;  Gaps    8;

Qy      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60

Qy     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagcc 120

Qy    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180

Qy    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 240

Qy    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300

Qy    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360

Qy    361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
          |||||||||||||||||||||||||| |||||||||| || |||||| |||| |||||
Db    361 aacgttgattttggtgcggaaggaagaaatcccgatcaacatcgttttacttataaaatt 420

Qy    421 gtgaaaagaaataattataagcctgacaattca---caccttacaacggcgattancat 477
          |||||| | |||||||||||  | |  | |  |  || |||||  ||||||||| |||
Db    421 gtgaaacggaataattataaagcagggactaaaggccatccttatggcggcgattatcat 480

Qy    478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
          |||||||||||||||||||||||||||||||||||||||||| |||||||| ||| | |||
Db    481 atgccgcgtttgcataaatttgtcacagatgcagaacctgttgaaatgaccagttatatg 540

Qy    538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
             |||   |  |||  |||| || | || || || || || ||||| || || ||  ||||
Db    541 gatgggcggaaatatatcgatcaaaataattaccctgaccgtgttcgtattggggcaggc 600

Qy    598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
             || |||||||| | |||||| ||       |||  || ||||    || ||
Db    601 aggcaatattggcgatctgatgaagatgagcccaataaccgcgaaagttcatatcatatt 660

Qy    652 gca----------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
          |||           ||| |  |||| ||||||||   |   ||  | |   |  |  |
Db    661 gcaagtgcgtattcttggctcgttggtggcaatacctttgcacaaaatggatcaggtggt 720

Qy    700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
          |||  | | |  ||  |  | ||         |    | |  |||||   | | || |
Db    721 ggcacagtcaacttaggtagtgaaaaaattaaacatagcccatatggtttttaccaaca 780

Qy    760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
          | |||  |   ||||||||  || || ||||||||||| |||||      |||  | ||
Db    781 ggaggctcatttggcgacagtggctcaccaatgtttatctatgatgcccaaaagcaaaag 840

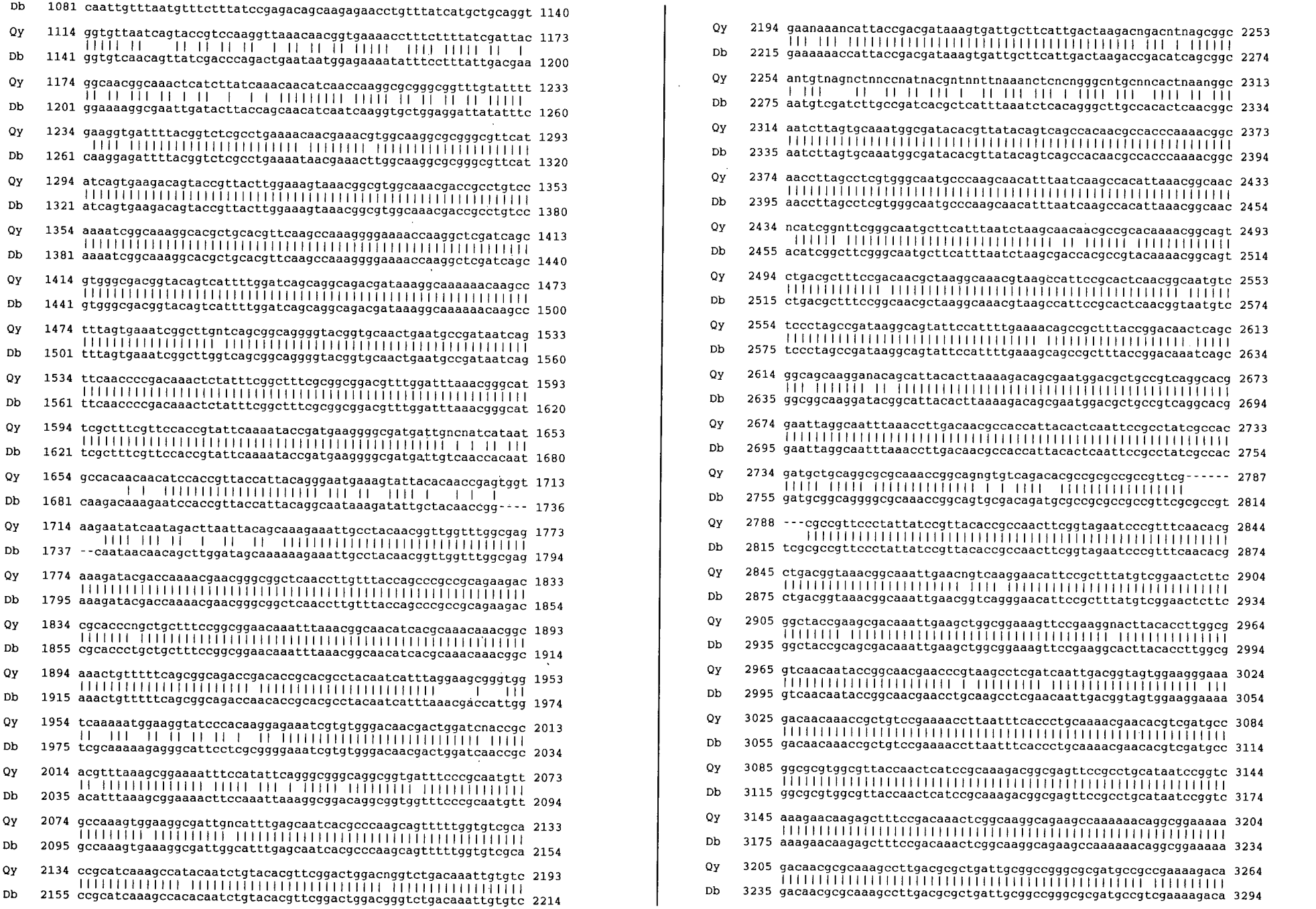
Qy    820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
          ||| |  | || || || || ||||| ||| |||| |||   || |     || ||| |||
Db    841 tggttaattaatggggtattgcaaacgggcaacccctatataggaaaaagcaatggcttc 900

Qy    880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
          |||||| | || |||||||||||||| ||||| || |     || ||||| ||| | ||
Db    901 cagctggttcgtaaagattggttctatgatgaaatctttgctggagatacccattcagta 960

Qy    940 tnttttgaaccgcgcagtaacggacatttttccttacatccaacaacaacggtacgggt 999
          |  |  ||||| ||    || ||  | |  || ||||    | | || || || || ||
Db    961 ttctacgaaccacgtcaaaatgggaaatactctttttaacgacgataataatggcacagga 1020

Qy   1000 acggtaacagaaaccaacgaaaaggtntccaatccaaa---gcttaaagtacagacagtc 1056
          |   | | |  |  |  | ||| |    ||    || ||     | |||  ||  || ||
Db   1021 aaaatcaatgccaaacatgaacacaattctctgcctaatagattaaaaacacgaaccgtt 1080

Qy   1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
          | | ||||| | |  |||||    || || |  | |||||| ||||   | || ||||| 
```



```
Qy   3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   3295 gaaagcgttgccgaaccggcccggcaggcaggcggggaaaatgtcggcattatgcaggcg 3354

Qy   3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
          |||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||
Db   3355 gaggaagagaaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaa 3414

Qy   3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
          |||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||
Db   3415 gcggaaacccggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttg 3474

Qy   3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
          ||||||| ||| |||||||||      || ||||||||||||||||||||| ||||||
Db   3475 ccgcaactgcaaccccaaccg------cagccccaaccgcagcgcgacctgatcagccgt 3528

Qy   3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3529 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3588

Qy   3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
          |||||||| |||||||| ||||||||||||||||||||| ||||||||||||.|||||
Db   3589 gacgaattagaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatc 3648

Qy   3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3649 cgggacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3708

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3709 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3768

Qy   3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3769 aaccggaccgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggc 3828

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
          |||||||||||||||||||||||| ||||||| |||||||||||||| ||||||||||||
Db   3829 gccgttttcgggcaatacggcatcgacaggttctacatcggcatcagcgcgggcgcgggt 3888

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
          |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||
Db   3889 tttagcagcggcagcctttcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3948

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   3949 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatc 4008

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4068

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
          |||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db   4069 acccccggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaa 4128

Qy   4105 ccggcgcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcg 4164
          ||||||||||||||| |||||||| |||||||| ||||||||||||||||||||||||||
Db   4129 ccggcgcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcg 4188

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   4189 ggcaaagtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgc 4248

Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   4249 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgcc 4308

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 gccgccaaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4368

Qy   4345 tggtaa 4350
          ||||||
Db   4369 tggtaa 4374
```

RESULT 6
AAA81477
ID AAA81477 standard; DNA; 26778 BP.
XX
AC AAA81477;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_25 SEQ ID NO:25.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 524-531; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 26778 BP; 6602 A; 7008 C; 6846 G; 6318 T; 4 other;

Query Match 80.6%; Score 3508; DB 21; Length 26778;

Best Local Similarity 88.4%; Pred. No. 0;
Matches 3876; Conservative 0; Mismatches 462; Indels 48; Gaps 8;

```
Qy       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
Db    9845 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 9904

Qy      61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
Db    9905 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagcc 9964

Qy     121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgacttttgccgaaaat 180
Db    9965 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 10024

Qy     181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
Db   10025 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 10084

Qy     241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
Db   10085 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 10144

Qy     301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
Db   10145 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 10204

Qy     361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
Db   10205 aacgttgattttggtgcggaaggaagaaatcccgatcaacatcgtttacttataaaatt 10264

Qy     421 gtgaaaagaaataattataagcctgacaattca---caccttacaacggcgattancat 477
Db   10265 gtgaaacggaataattataaagcagggactaaaggccatccttatggcggcgattatcat 10324

Qy     478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
Db   10325 atgccgcgtttgcataaatttgtcacagatgcagaacctgttgaaatgaccagttatatg 10384

Qy     538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
Db   10385 gatgggcggaaatatatcgatcaaaataattaccctgaccgtgttcgtattggggcaggc 10444

Qy     598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
Db   10445 aggcaatattggcgatctgatgaagatgagcccaataaccgcgaaagttcatatcatatt 10504

Qy     652 gca----------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
Db   10505 gcaagtgcgtattcttggctcgttggtggcaatacctttgcacaaaatggatcaggtggt 10564

Qy     700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
Db   10565 ggcacagtcaacttaggtagtgaaaaaattaaacatagcccatatggtttttaccaaca 10624

Qy     760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
Db   10625 ggaggctcatttggcgacagtggctcaccaatgtttatctatgatgcccaaaagcaaaag 10684

Qy     820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
Db   10685 tggttaattaatggggtattgcaaacgggcaacccctatataggaaaaagcaatggcttc 10744

Qy     880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
Db   10745 cagctggttcgtaaagattggttctatgatgaaatctttgctggagatacccattcagta 10804

Qy     940 tnttttgaaccgcgcagtaacggacatttttcctttacatccaacaacaacggtacgggt 999
Db   10805 ttctacgaaccacgtcaaaatgggaaatactcttttaacgacgataataatggcacagga 10864

Qy    1000 acggtaacagaaaccaacgaaaaggtntccaatccaaa---gcttaaagtacagacagtc 1056
Db   10865 aaaatcaatgccaaacatgaacacaattctctgcctaatagattaaaaacacgaaccgtt 10924

Qy    1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
Db   10925 caattgtttaatgtttctttatccgagacagcaagagaacctgtttatcatgctgcaggt 10984

Qy    1114 ggtgttaatcagtaccgtccaaggttaaacaacggtgaaaacctttctttttatcgattac 1173
Db   10985 ggtgtcaacagttatcgacccagactgaataatggagaaaatatttcctttattgacgaa 11044

Qy    1174 ggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttt 1233
Db   11045 ggaaaaggcgaattgatacttaccagcaacatcaatcaaggtgctggaggattatatttc 11104

Qy    1234 gaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcat 1293
Db   11105 caaggagattttacggtctcgcctgaaaataacgaaacttggcaaggcgcgggcgttcat 11164

Qy    1294 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1353
Db   11165 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 11224

Qy    1354 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1413
Db   11225 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 11284

Qy    1414 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1473
Db   11285 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 11344

Qy    1474 tttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcag 1533
Db   11345 tttagtgaaatcggcttggtcagcggcaggggtacggtgcaactgaatgccgataatcag 11404

Qy    1534 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1593
Db   11405 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 11464

Qy    1594 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcataat 1653
Db   11465 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgtcaaccacaat 11524

Qy    1654 gccacaacaacatccaccgttaccattacagggaatgaaagtattacacaaccgagtggt 1713
Db   11525 caagacaaagaatccaccgttaccattacaggcaataaagatattgctacaaccgg---- 11580

Qy    1714 aagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtttggcgag 1773
Db   11581 --caataacaacagcttggatagcaaaaaagaaattgcctacaacggttggtttggcgag 11638

Qy    1774 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1833
Db   11639 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 11698

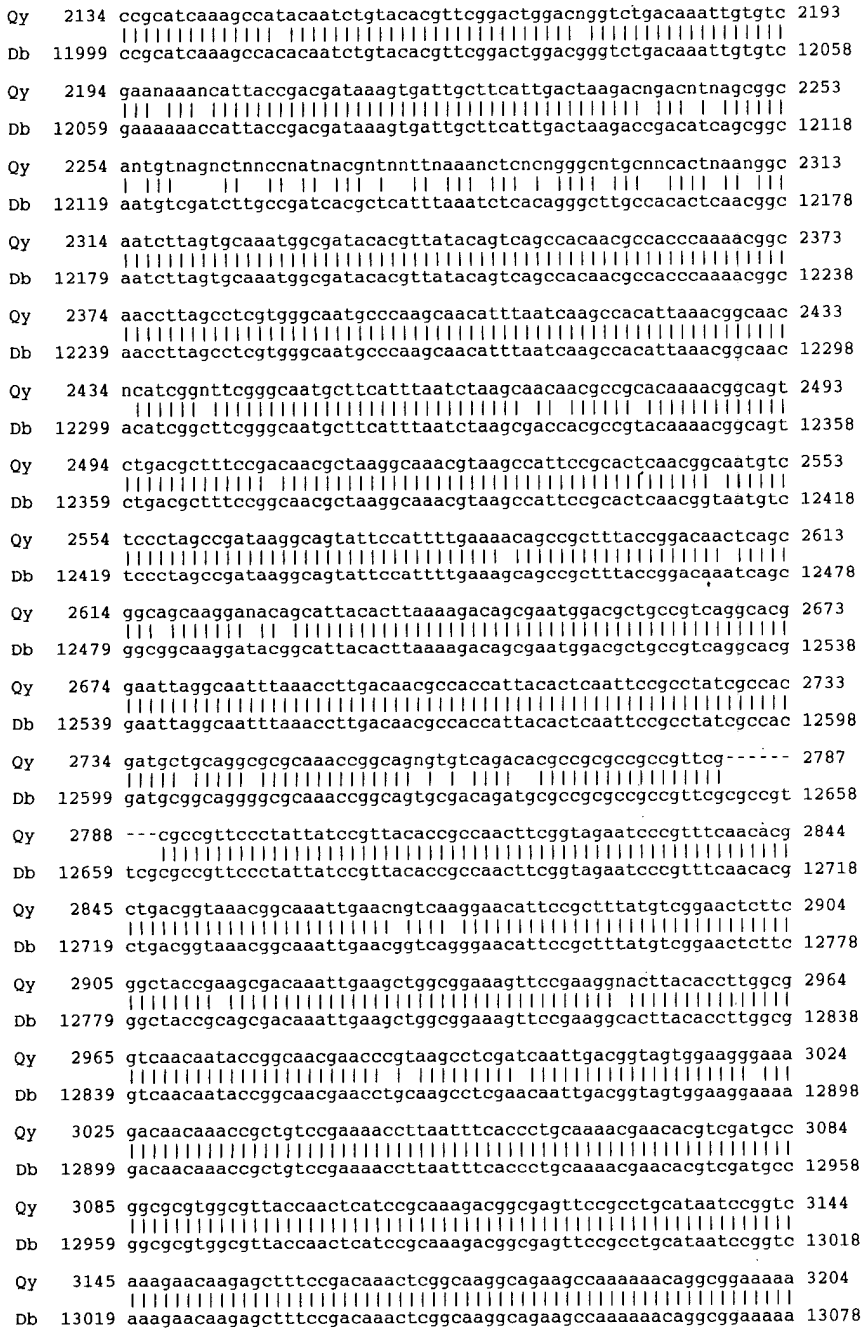
Qy    1834 cgcacccngctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1893
Db   11699 cgcaccctgctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 11758

Qy    1894 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1953
Db   11759 aaactgtttttcagcggcagaccaacaccgcacgcctacaatcatttaaacgaccattgg 11818

Qy    1954 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgactggatcnaccgc 2013
Db   11819 tcgcaaaaagagggcattcctcgcgggggaaatcgtgtgggacaacgactggatcaaccgc 11878

Qy    2014 acgtttaaagcggaaaatttccatattcagggcgggcaggcggtgatttcccgcaatgtt 2073
Db   11879 acatttaaagcggaaaacttccaaattaaaggcggacaggcggtggtttcccgcaatgtt 11938

Qy    2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagtttttggtgtcgca 2133
Db   11939 gccaaagtgaaaggcgattggcatttgagcaatcacgcccaagcagtttttggtgtcgca 11998
```


Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193
Db 11999 ccgcatcaaagccacacaatctgtacacgttcggactggacgggtctgacaaattgtgtc 12058
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 12059 gaaaaaaccattaccgacgataaagtgattgcttcattgactaagaccgacatcagcggc 12118
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 12119 aatgtcgatcttgccgatcacgctcatttaaatctcacagggcttgccacactcaacggc 12178
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 12179 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 12238
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db 12239 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 12298
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 12299 acatcggcttcgggcaatgcttcatttaatctaagcgaccacgccgtacaaaacggcagt 12358
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 12359 ctgacgctttccggcaacgctaaggcaaacgtaagccattccgcactcaacggtaatgtc 12418
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 12419 tccctagccgataaggcagtattccattttgaaagcagccgctttaccggacaaatcagc 12478
Qy 2614 ggcagcaaggcanacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 12479 ggcggcaaggatacggcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 12538
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 12539 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 12598
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcg------ 2787
Db 12599 gatgcggcaggggcgcaaaccggcagtgcgacagatgcgccgcgccgccgttcgcgccgt 12658
Qy 2788 ---cgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2844
Db 12659 tcgcgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 12718
Qy 2845 ctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttc 2904
Db 12719 ctgacggtaaacggcaaattgaacggtcagggaacattccgctttatgtcggaactcttc 12778
Qy 2905 ggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcg 2964
Db 12779 ggctaccgcagcgacaaattgaagctggcggaaagttccgaaggcacttacaccttggcg 12838
Qy 2965 gtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaa 3024
Db 12839 gtcaacaataccggcaacgaacctgcaagcctcgaacaattgacggtagtggaagggaaa 12898
Qy 3025 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3084
Db 12899 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 12958
Qy 3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
Db 12959 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 13018
Qy 3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204
Db 13019 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 13078
Qy 3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
Db 13079 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgtcgaaaagaca 13138
Qy 3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
Db 13139 gaaagcgttgccgaaccggcccggcaggcaggcggggaaaatgtcggcattatgcaggcg 13198
Qy 3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
Db 13199 gaggaagagaaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaa 13258
Qy 3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
Db 13259 gcggaaacccggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttg 13318
Qy 3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
Db 13319 ccgcaactgcaaccccaaccg------cagccccaaccgcagcgcgacctgatcagccgt 13372
Qy 3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
Db 13373 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 13432
Qy 3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
Db 13433 gacgaattagaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatc 13492
Qy 3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
Db 13493 cgggacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 13552
Qy 3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
Db 13553 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 13612
Qy 3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
Db 13613 aaccggaccgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggc 13672
Qy 3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
Db 13673 gccgttttcgggcaatacggcatcgacaggttctacatcggcatcagcgcgggcgcgggt 13732
Qy 3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
Db 13733 tttagcagcggcagcctttcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 13792
Qy 3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
Db 13793 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatc 13852
Qy 3985 ggcgcaacgcgctatttcgtccaaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
Db 13853 ggcgcaacgcgctatttcgtccaaaaaagcggattaccgctacgaaaacgtcaatatcgcc 13912
Qy 4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
Db 13913 acccccggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaa 13972
Qy 4105 ccggcgcaacacatntccatcacnccttatttnagcctgtcctataccgatgccgcttcg 4164
Db 13973 ccggcgcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcg 14032
Qy 4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
Db 14033 ggcaaagtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgc 14092
Qy 4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
Db 14093 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgcc 14152
Qy 4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344

```
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14153 gccgccaaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 14212

Qy   4345 tggtaa 4350
          ||||||
Db  14213 tggtaa 14218


RESULT   7
AAF21612/c
ID   AAF21612 standard; DNA; 349980 BP.
XX
AC   AAF21612;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;

  Query Match             80.6%;  Score 3508;  DB 21;  Length 349980;
  Best Local Similarity   88.4%;  Pred. No. 0;
  Matches 3876;  Conservative    0;  Mismatches  462;  Indels   48;  Gaps    8;

Qy       1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  289191 ATGAAAACAACCGACAAACGGACAACCGAAACACACCGCAAAGCCCCGAAAACCGGCCGC 289132

Qy      61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  289131 ATCCGCTTCTCGCCTGCTTACTTAGCCATATGCCTGTCGTTCGGCATTCTTCCCCAAGCC 289072

Qy     121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  289071 TGGGCGGGACACACTTATTTCGGCATCAACTACCAATACTATCGCGACTTTGCCGAAAAT 289012

Qy     181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  289011 AAAGGCAAGTTTGCAGTCGGGGCGAAAGATATTGAGGTTTACAACAAAAAAGGGGAGTTG 288952

Qy     241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  288951 GTCGGCAAATCAATGACAAAAGCCCCGATGATTGATTTTTCTGTGGTGTCGCGTAACGGC 288892

Qy     301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  288891 GTGGCGGCATTGGTGGGCGATCAATATATTGTGAGCGTGGCACATAACGGCGGCTATAAC 288832

Qy     361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
           |||||||||||||||||||||||||| |||||||||| || |||||| |||  |||||
Db  288831 AACGTTGATTTTGGTGCGGAAGGAAGAAATCCCGATCAACATCGTTTTACTTATAAAATT 288772

Qy     421 gtgaaaagaaataattataagcctgacaattca---caccctacaacggcgattancat 477
           |||||| | |||||||||||  | |  | |  |   || |||||   |||||||||| |||
Db  288771 GTGAAACGGAATAATTATAAAGCAGGGACTAAAGGCCATCCTTATGGCGGCGATTATCAT 288712

Qy     478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||| | |||
Db  288711 ATGCCGCGTTTGCATAAATTTGTCACAGATGCAGAACCTGTTGAAATGACCAGTTATATG 288652

Qy     538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
             |||   |  |||  |||| || | || || || || ||||| || || ||  ||||
Db  288651 GATGGGCGGAAATATATCGATCAAAATAATTACCCTGACCGTGTTCGTATTGGGGCAGGC 288592

Qy     598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
             || |||||||| | |||||| ||       |||  || |||    || ||
Db  288591 AGGCAATATTGGCGATCTGATGAAGATGAGCCCAATAACCGCGAAAGTTCATATCATATT 288532

Qy     652 gca------------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
           |||             ||| |  |||| ||||||||   |   ||  | |   |  |  |
Db  288531 GCAAGTGCGTATTCTTGGCTCGTTGGTGGCAATACCTTTGCACAAAATGGATCAGGTGGT 288472

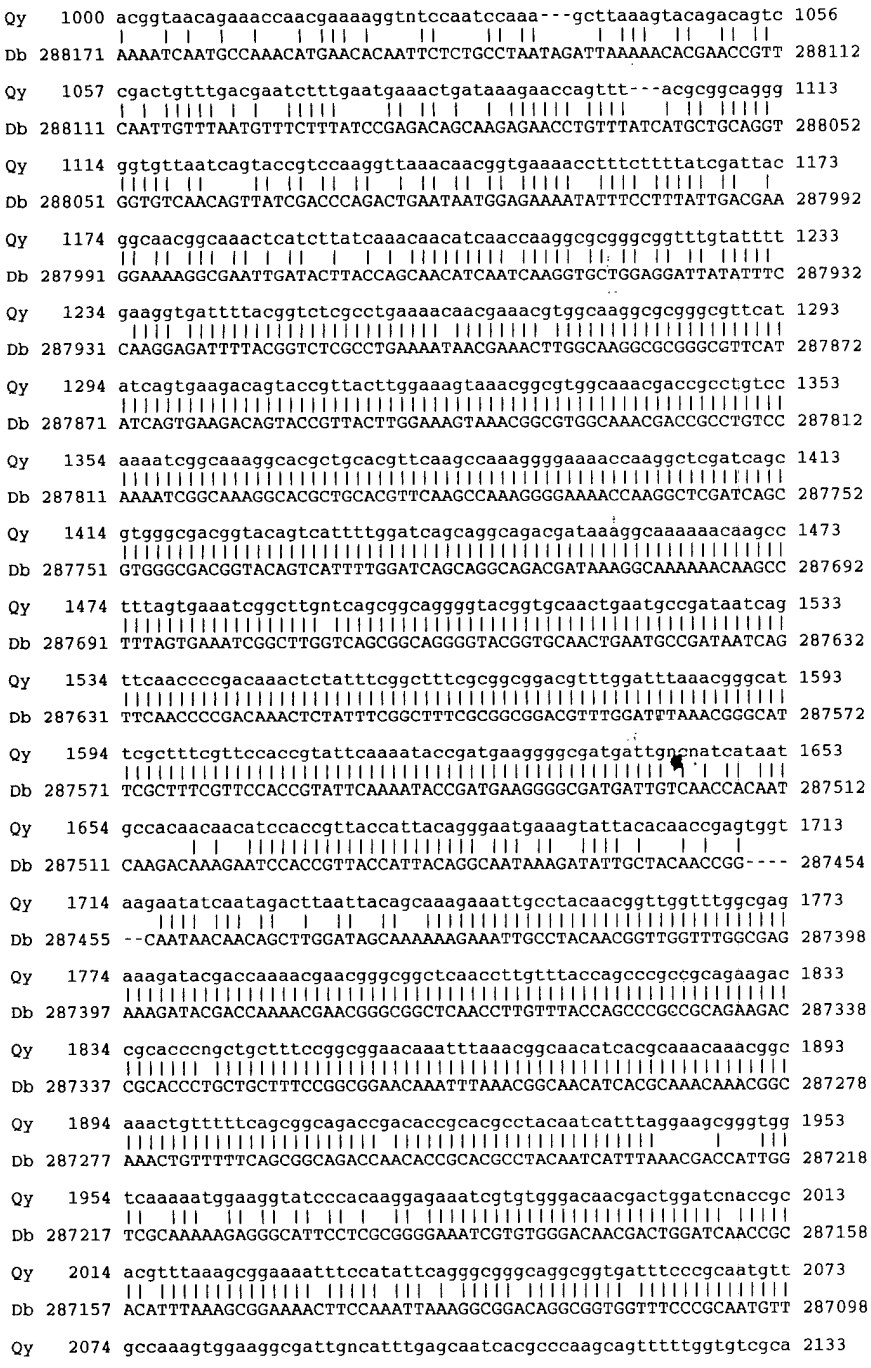
Qy     700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
           |||  | | |  ||  |  | ||       |    | |   |||||   | | || |
Db  288471 GGCACAGTCAACTTAGGTAGTGAAAAAATTAAACATAGCCCATATGGTTTTTTACCAACA 288412

Qy     760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
           | |||  |    |||||||| || || ||||||||||| |||||      |||  | ||
Db  288411 GGAGGCTCATTTGGCGACAGTGGCTCACCAATGTTTATCTATGATGCCCAAAAGCAAAAG 288352

Qy     820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
           ||| |  | || || || || ||||| ||| |||| |||   || |     || || |||
Db  288351 TGGTTAATTAATGGGGTATTGCAAACGGGCAACCCCTATATAGGAAAAAGCAATGGCTTC 288292

Qy     880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
           |||||| | || |||||||||||||| ||||| || |     || ||||| ||| | ||
Db  288291 CAGCTGGTTCGTAAAGATTGGTTCTATGATGAAATCTTTGCTGGAGATACCCATTCAGTA 288232

Qy     940 tnttttgaaccgcgcagtaacggacatttttccttacatccaacaacaacggtacgggt 999
           |  |  ||||| ||    || ||  | |  || |||     | | || || || || ||
Db  288231 TTCTACGAACCACGTCAAAATGGGAAATACTCTTTTAACGACGATAATAATGGCACAGGA 288172
```


Qy 1000 acggtaacagaaaccaacgaaaaggtntccaatccaaa---gcttaaagtacagacagtc 1056
Db 288171 AAAATCAATGCCAAACATGAACACAATTCTCTGCCTAATAGATTAAAAACACGAACCGTT 288112
Qy 1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
Db 288111 CAATTGTTTAATGTTTCTTTATCCGAGACAGCAAGAGAACCTGTTTATCATGCTGCAGGT 288052
Qy 1114 ggtgttaatcagtaccgtccaaggttaaacaacggtgaaaaccttttcttttatcgattac 1173
Db 288051 GGTGTCAACAGTTATCGACCCAGACTGAATAATGGAGAAAATATTTCCTTTATTGACGAA 287992
Qy 1174 ggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttt 1233
Db 287991 GGAAAAGGCGAATTGATACTTACCAGCAACATCAATCAAGGTGCTGGAGGATTATATTTC 287932
Qy 1234 gaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcat 1293
Db 287931 CAAGGAGATTTTACGGTCTCGCCTGAAAATAACGAAACTTGGCAAGGCGCGGGCGTTCAT 287872
Qy 1294 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1353
Db 287871 ATCAGTGAAGACAGTACCGTTACTTGGAAAGTAAACGGCGTGGCAAACGACCGCCTGTCC 287812
Qy 1354 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1413
Db 287811 AAAATCGGCAAAGGCACGCTGCACGTTCAAGCCAAAGGGGAAAACCAAGGCTCGATCAGC 287752
Qy 1414 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1473
Db 287751 GTGGGCGACGGTACAGTCATTTTGGATCAGCAGGCAGACGATAAAGGCAAAAAACAAGCC 287692
Qy 1474 tttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcag 1533
Db 287691 TTTAGTGAAATCGGCTTGGTCAGCGGCAGGGGTACGGTGCAACTGAATGCCGATAATCAG 287632
Qy 1534 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1593
Db 287631 TTCAACCCCGACAAACTCTATTTCGGCTTTCGCGGCGGACGTTTGGATTTAAACGGGCAT 287572
Qy 1594 tcgctttcgttccaccgtattcaaaatacCgatgaaggggcgatgattgngnatcataat 1653
Db 287571 TCGCTTTCGTTCCACCGTATTCAAAATACCGATGAAGGGGCGATGATTGTCAACCACAAT 287512
Qy 1654 gccacaacaacatccaccgttaccattacagggaatgaaagtattacacaaccgagtggt 1713
Db 287511 CAAGACAAAGAATCCACCGTTACCATTACAGGCAATAAAGATATTGCTACAACCGG---- 287454
Qy 1714 aagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtttggcgag 1773
Db 287455 --CAATAACAACAGCTTGGATAGCAAAAAAGAAATTGCCTACAACGGTTGGTTTGGCGAG 287398
Qy 1774 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1833
Db 287397 AAAGATACGACCAAAACGAACGGGCGGCTCAACCTTGTTTACCAGCCCGCCGCAGAAGAC 287338
Qy 1834 cgcacccngctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1893
Db 287337 CGCACCCTGCTGCTTTCCGGCGGAACAAATTTAAACGGCAACATCACGCAAACAAACGGC 287278
Qy 1894 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1953
Db 287277 AAACTGTTTTTCAGCGGCAGACCAACACCGCACGCCTACAATCATTTAAACGACCATTGG 287218
Qy 1954 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgactggatcnaccgc 2013
Db 287217 TCGCAAAAAGAGGGCATTCCTCGCGGGGAAATCGTGTGGGACAACGACTGGATCAACCGC 287158
Qy 2014 acgtttaaagcggaaaatttccatattcagggcgggcaggcggtgatttcccgcaatgtt 2073
Db 287157 ACATTTAAAGCGGAAAACTTCCAAATTAAAGGCGGACAGGCGGTGGTTTCCCGCAATGTT 287098
Qy 2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagtttttggtgtcgca 2133
Db 287097 GCCAAAGTGAAAGGCGATTGGCATTTGAGCAATCACGCCCAAGCAGTTTTTGGTGTCGCA 287038
Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193
Db 287037 CCGCATCAAAGCCACACAATCTGTACACGTTCGGACTGGACGGGTCTGACAAATTGTGTC 286978
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 286977 GAAAAAACCATTACCGACGATAAAGTGATTGCTTCATTGACTAAGACCGACATCAGCGGC 286918
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 286917 AATGTCGATCTTGCCGATCACGCTCATTTAAATCTCACAGGGCTTGCCACACTCAACGGC 286858
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 286857 AATCTTAGTGCAAATGGCGATACACGTTATACAGTCAGCCACAACGCCACCCAAAACGGC 286798
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db 286797 AACCTTAGCCTCGTGGGCAATGCCCAAGCAACATTTAATCAAGCCACATTAAACGGCAAC 286738
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 286737 ACATCGGCTTCGGGCAATGCTTCATTTAATCTAAGCGACCACGCCGTACAAAACGGCAGT 286678
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 286677 CTGACGCTTTCCGGCAACGCTAAGGCAAACGTAAGCCATTCCGCACTCAACGGTAATGTC 286618
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 286617 TCCCTAGCCGATAAGGCAGTATTCCATTTTGAAAGCAGCCGCTTTACCGGACAAATCAGC 286558
Qy 2614 ggcagcaaggganacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 286557 GGCGGCAAGGATACGGCATTACACTTAAAAGACAGCGAATGGACGCTGCCGTCAGGCACG 286498
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 286497 GAATTAGGCAATTTAAACCTTGACAACGCCACCATTACACTCAATTCCGCCTATCGCCAC 286438
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcg------ 2787
Db 286437 GATGCGGCAGGGGCGCAAACCGGCAGTGCGACAGATGCGCCGCGCCGCCGTTCGCGCCGT 286378
Qy 2788 ---cgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2844
Db 286377 TCGCGCCGTTCCCTATTATCCGTTACACCGCCAACTTCGGTAGAATCCCGTTTCAACACG 286318
Qy 2845 ctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttc 2904
Db 286317 CTGACGGTAAACGGCAAATTGAACGGTCAGGGAACATTCCGCTTTATGTCGGAACTCTTC 286258
Qy 2905 ggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcg 2964
Db 286257 GGCTACCGCAGCGACAAATTGAAGCTGGCGGAAAGTTCCGAAGGCACTTACACCTTGGCG 286198
Qy 2965 gtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaa 3024
Db 286197 GTCAACAATACCGGCAACGAACCTGCAAGCCTCGAACAATTGACGGTAGTGGAAGGAAAA 286138
Qy 3025 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3084
Db 286137 GACAACAAACCGCTGTCCGAAAACCTTAATTTCACCCTGCAAAACGAACACGTCGATGCC 286078
Qy 3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
Db 286077 GGCGCGTGGCGTTACCAACTCATCCGCAAAGACGGCGAGTTCCGCCTGCATAATCCGGTC 286018
Qy 3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204

```
Db 286017 AAAGAACAAGAGCTTTCCGACAAACTCGGCAAGGCAGAAGCCAAAAAACAGGCGGAAAAA 285958

Qy   3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 285957 GACAACGCGCAAAGCCTTGACGCGCTGATTGCGGCCGGGCGCGATGCCGTCGAAAAGACA 285898

Qy   3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 285897 GAAAGCGTTGCCGAACCGGCCCGGCAGGCAGGCGGGGAAAATGTCGGCATTATGCAGGCG 285838

Qy   3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
          |||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||
Db 285837 GAGGAAGAGAAAAAACGGGTGCAGGCGGATAAAGACACCGCCTTGGCGAAACAGCGCGAA 285778

Qy   3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
          |||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||||||
Db 285777 GCGGAAACCCGGCCGGCTACCACCGCCTTCCCCCGCGCCCGCCGCGCCCGCCGGGATTTG 285718

Qy   3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
          ||||||| ||| |||||||||      || |||||||||||||||||||||| ||||||
Db 285717 CCGCAACTGCAACCCCAACCG------CAGCCCCAACCGCAGCGCGACCTGATCAGCCGT 285664

Qy   3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 285663 TATGCCAATAGCGGTTTGAGTGAATTTTCCGCCACGCTCAACAGCGTTTTCGCCGTACAG 285604

Qy   3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
          |||||||| |||||||| |||||||||||||||||||| |||||||||||| |||||
Db 285603 GACGAATTAGACCGCGTATTTGCCGAAGACCGCCGCAACGCCGTTTGGACAAGCGGCATC 285544

Qy   3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 285543 CGGGACACCAAACACTACCGTTCGCAAGATTTCCGCGCCTACCGCCAACAAACCGACCTG 285484

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 285483 CGCCAAATCGGTATGCAGAAAAACCTCGGCAGCGGGCGCGTCGGCATCCTGTTTTCGCAC 285424

Qy   3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 285423 AACCGGACCGAAAACACCTTCGACGACGGCATCGGCAACTCGGCACGGCTTGCCCACGGC 285364

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
          |||||||||||||||||||||||| ||||||| |||||||||||| ||||||||||
Db 285363 GCCGTTTTCGGGCAATACGGCATCGACAGGTTCTACATCGGCATCAGCGCGGGCGCGGGT 285304

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
          |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||
Db 285303 TTTAGCAGCGGCAGCCTTTCAGACGGCATCGGAGGCAAAATCCGCCGCCGCGTGCTGCAT 285244

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db 285243 TACGGCATTCAGGCACGATACCGCGCCGGTTTCGGCGGATTCGGCATCGAACCGCACATC 285184

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 285183 GGCGCAACGCGCTATTTCGTCCAAAAAGCGGATTACCGCTACGAAAACGTCAATATCGCC 285124

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
          |||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||
Db 285123 ACCCCCGGCCTTGCATTCAACCGCTACCGCGCGGGCATTAAGGCAGATTATTCATTCAAA 285064

Qy   4105 ccggcgcaacacatntccatcacnccttatttnagcctgtcctataccgatgccgcttcg 4164
          |||||||||||||| |||||||| |||||||| ||||||||||||||||||||||||||||
Db 285063 CCGGCGCAACACATTTCCATCACGCCTTATTTGAGCCTGTCCTATACCGATGCCGCTTCG 285004

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 285003 GGCAAAGTCCGAACACGCGTCAATACCGCCGTATTGGCTCAGGATTTCGGCAAAACCCGC 284944

Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db 284943 AGTGCGGAATGGGGCGTAAACGCCGAAATCAAAGGTTTCACGCTGTCCCTCCACGCTGCC 284884

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db 284883 GCCGCCAAAGGCCCGCAACTGGAAGCGCAACACAGCGCGGGCATCAAATTAGGCTACCGC 284824

Qy   4345 tggtaa 4350
          ||||||
Db 284823 TGGTAA 284818
```

RESULT 8
AAA81489/c
ID AAA81489 standard; DNA; 837096 BP.
XX
AC AAA81489;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 629-865; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than

CC other more variable regions.
XX
SQ Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;

Query Match 80.6%; Score 3508; DB 21; Length 837096;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 3876; Conservative 0; Mismatches 462; Indels 48; Gaps 8;

```
Qy      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653962 ATGAAAACAACCGACAAACGGACAACCGAAACACACCGCAAAGCCCCGAAAACCGGCCGC 653903

Qy     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653902 ATCCGCTTCTCGCCTGCTTACTTAGCCATATGCCTGTCGTTCGGCATTCTTCCCCAAGCC 653843

Qy    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653842 TGGGCGGGACACACTTATTTCGGCATCAACTACCAATACTATCGCGACTTTGCCGAAAAT 653783

Qy    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 653782 AAAGGCAAGTTTGCAGTCGGGGCGAAAGATATTGAGGTTTACAACAAAAAAGGGGAGTTG 653723

Qy    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653722 GTCGGCAAATCAATGACAAAAGCCCCGATGATTGATTTTTCTGTGGTGTCGCGTAACGGC 653663

Qy    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 653662 GTGGCGGCATTGGTGGGCGATCAATATATTGTGAGCGTGGCACATAACGGCGGCTATAAC 653603

Qy    361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
          |||||||||||||||||||||||||||||| |||||||||| || |||||| |||| |||||
Db 653602 AACGTTGATTTTGGTGCGGAAGGAAGAAATCCCGATCAACATCGTTTTACTTATAAAATT 653543

Qy    421 gtgaaaagaaataattatataagcctgacaattca---caccctacaacggcgattancat 477
          |||||| | |||||||||| | | | | || |||||  |||||||| |||
Db 653542 GTGAAACGGAATAATTATAAAGCAGGGACTAAAGGCCATCCTTATGGCGGCGATTATCAT 653483

Qy    478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
          |||||||||||||||||||||||||||||||||||||||| |||||||| ||| | |||
Db 653482 ATGCCGCGTTTGCATAAATTTGTCACAGATGCAGAACCTGTTGAAATGACCAGTTATATG 653423

Qy    538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
             |||  |  |||  |||| || | || || || |||| || || ||  ||||
Db 653422 GATGGGCGGAAATATATCGATCAAAATAATTACCCTGACCGTGTTCGTATTGGGGCAGGC 653363

Qy    598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
             || |||||||| | |||||| ||      |||  || ||||    || ||
Db 653362 AGGCAATATTGGCGATCTGATGAAGATGAGCCCAATAACCGCGAAAGTTCATATCATATT 653303

Qy    652 gca------------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
          |||             ||| |  |||| ||||||||   |   ||  | |  |  |  |
Db 653302 GCAAGTGCGTATTCTTGGCTCGTTGGTGGCAATACCTTTGCACAAAATGGATCAGGTGGT 653243

Qy    700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
          |||  | | |  ||  |  | ||        |    | |  |||||   | | || |
Db 653242 GGCACAGTCAACTTAGGTAGTGAAAAAATTAAACATAGCCCATATGGTTTTTTACCAACA 653183

Qy    760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
          | |||  |    |||||||| || || |||||||||| |||||      |||  | ||
Db 653182 GGAGGCTCATTTGGCGACAGTGGCTCACCAATGTTTATCTATGATGCCCAAAAGCAAAAG 653123

Qy    820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
          ||| |  | || || || || ||||| ||| |||| |||   || |     || || |||
Db 653122 TGGTTAATTAATGGGGTATTGCAAACGGGCAACCCCTATATAGGAAAAAGCAATGGCTTC 653063

Qy    880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
          |||||| | || |||||||||||||| ||||| || |     || ||||| ||| | ||
Db 653062 CAGCTGGTTCGTAAAGATTGGTTCTATGATGAAATCTTTGCTGGAGATACCCATTCAGTA 653003

Qy    940 tnttttgaaccgcgcagtaacggacatttttcctttacatccaacaacaacggtacgggt 999
          |  | ||||| ||    || ||  | |  || ||||    | | || || || || ||
Db 653002 TTCTACGAACCACGTCAAAATGGGAAATACTCTTTTAACGACGATAATAATGGCACAGGA 652943

Qy   1000 acggtaacagaaaccaacgaaaaggtntccaatccaaa---gcttaaagtacagacagtc 1056
          |  | |  |  |  | ||| |   ||   || ||    | |||  ||  || ||
Db 652942 AAAATCAATGCCAAACATGAACACAATTCTCTGCCTAATAGATTAAAAACACGAACCGTT 652883

Qy   1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
          | | ||||| | |  |||||    || || |  | ||||||  ||||   | || |||||
Db 652882 CAATTGTTTAATGTTTCTTTATCCGAGACAGCAAGAGAACCTGTTTATCATGCTGCAGGT 652823

Qy   1114 ggtgttaatcagtaccgtccaaggttaaacaacggtgaaaacctttcttttttatcgattac 1173
          |||||  ||    || || || ||  | || ||  |||||  ||||  |||||  ||  |
Db 652822 GGTGTCAACAGTTATCGACCCAGACTGAATAATGGAGAAAATATTTCCTTTATTGACGAA 652763

Qy   1174 ggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttt 1233
          || || ||| || | ||  |  | | |||||||||  |||||  || || || || |||||
Db 652762 GGAAAAGGCGAATTGATACTTACCAGCAACATCAATCAAGGTGCTGGAGGATTATATTTC 652703

Qy   1234 gaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcat 1293
          ||||  |||||||||||||||||||||||||||||  ||||||||  ||||||||||||||||||||
Db 652702 CAAGGAGATTTTACGGTCTCGCCTGAAAATAACGAAACTTGGCAAGGCGCGGGCGTTCAT 652643

Qy   1294 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652642 ATCAGTGAAGACAGTACCGTTACTTGGAAAGTAAACGGCGTGGCAAACGACCGCCTGTCC 652583

Qy   1354 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652582 AAAATCGGCAAAGGCACGCTGCACGTTCAAGCCAAAGGGGAAAACCAAGGCTCGATCAGC 652523

Qy   1414 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652522 GTGGGCGACGGTACAGTCATTTTGGATCAGCAGGCAGACGATAAAGGCAAAAAACAAGCC 652463

Qy   1474 tttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcag 1533
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 652462 TTTAGTGAAATCGGCTTGGTCAGCGGCAGGGGTACGGTGCAACTGAATGCCGATAATCAG 652403

Qy   1534 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652402 TTCAACCCCGACAAACTCTATTTCGGCTTTCGCGGCGGACGTTTGGATTTAAACGGGCAT 652343

Qy   1594 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcataat 1653
          ||||||||||||||||||||||||||||||||||||||||||||||||||| | | || |||
Db 652342 TCGCTTTCGTTCCACCGTATTCAAAATACCGATGAAGGGGCGATGATTGTCAACCACAAT 652283

Qy   1654 gccacaacaacatccaccgttaccattacagggaatgaaagtattacacaaccgagtggt 1713
               | |  |||||||||||||||||||| ||| ||  |||| |   | |  |
Db 652282 CAAGACAAAGAATCCACCGTTACCATTACAGGCAATAAAGATATTGCTACAACCGG---- 652225

Qy   1714 aagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtttggcgag 1773
             |||| ||| ||  |  ||  ||  ||||||||||||||||||||||||||||||||
Db 652226 --CAATAACAACAGCTTGGATAGCAAAAAAGAAATTGCCTACAACGGTTGGTTTGGCGAG 652169

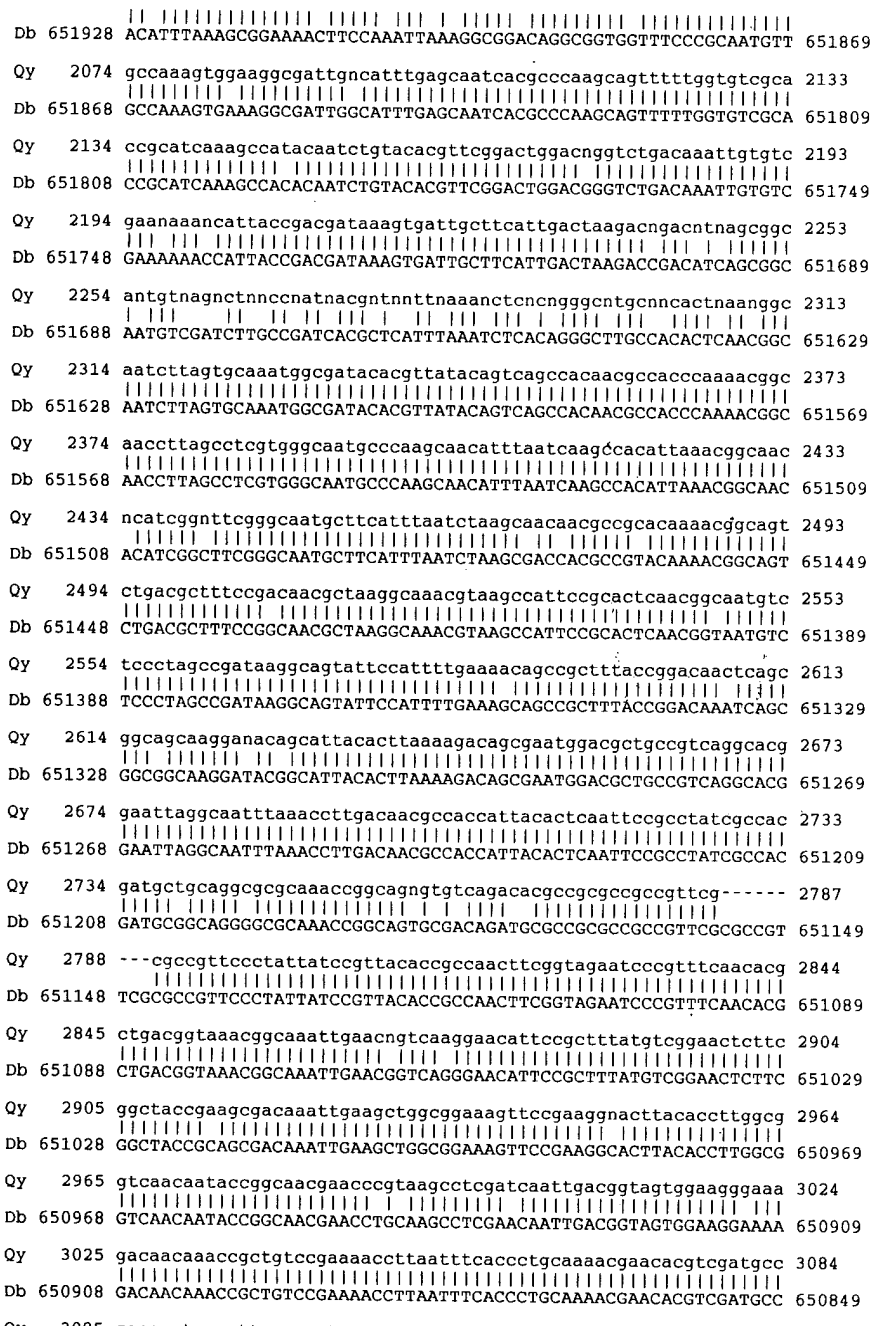
Qy   1774 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652168 AAAGATACGACCAAAACGAACGGGCGGCTCAACCTTGTTTACCAGCCCGCCGCAGAAGAC 652109

Qy   1834 cgcacccngctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1893
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 652108 CGCACCCTGCTGCTTTCCGGCGGAACAAATTTAAACGGCAACATCACGCAAACAAACGGC 652049

Qy   1894 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1953
          |||||||||||||||||||| ||||||||||||||||||||||||     |  |||
Db 652048 AAACTGTTTTTCAGCGGCAGACCAACACCGCACGCCTACAATCATTTAAACGACCATTGG 651989

Qy   1954 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgactggatcnaccgc 2013
          ||  |||  || || || || |  || ||||||||||||||||||||||||||| |||||
Db 651988 TCGCAAAAAGAGGGCATTCCTCGCGGGGAAATCGTGTGGGACAACGACTGGATCAACCGC 651929

Qy   2014 acgtttaaagcggaaaatttccatattcagggcgggcaggcggtgatttcccgcaatgtt 2073
```


Db 651928 ACATTTAAAGCGGAAAACTTCCAAATTAAAGGCGGACAGGCGGTGGTTTCCCGCAATGTT 651869
Qy 2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagtttttggtgtcgca 2133
Db 651868 GCCAAAGTGAAAGGCGATTGGCATTTGAGCAATCACGCCCAAGCAGTTTTTGGTGTCGCA 651809
Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193
Db 651808 CCGCATCAAAGCCACACAATCTGTACACGTTCGGACTGGACGGGTCTGACAAATTGTGTC 651749
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 651748 GAAAAAACCATTACCGACGATAAAGTGATTGCTTCATTGACTAAGACCGACATCAGCGGC 651689
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 651688 AATGTCGATCTTGCCGATCACGCTCATTTAAATCTCACAGGGCTTGCCACACTCAACGGC 651629
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 651628 AATCTTAGTGCAAATGGCGATACACGTTATACAGTCAGCCACAACGCCACCCAAAACGGC 651569
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagcacattaaacggcaac 2433
Db 651568 AACCTTAGCCTCGTGGGCAATGCCCAAGCAACATTTAATCAAGCCACATTAAACGGCAAC 651509
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 651508 ACATCGGCTTCGGGCAATGCTTCATTTAATCTAAGCGACCACGCCGTACAAAACGGCAGT 651449
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 651448 CTGACGCTTTCCGGCAACGCTAAGGCAAACGTAAGCCATTCCGCACTCAACGGTAATGTC 651389
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 651388 TCCCTAGCCGATAAGGCAGTATTCCATTTTGAAAGCAGCCGCTTTACCGGACAAATCAGC 651329
Qy 2614 ggcagcaaggganacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 651328 GGCGGCAAGGATACGGCATTACACTTAAAAGACAGCGAATGGACGCTGCCGTCAGGCACG 651269
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 651268 GAATTAGGCAATTTAAACCTTGACAACGCCACCATTACACTCAATTCCGCCTATCGCCAC 651209
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcg------ 2787
Db 651208 GATGCGGCAGGGGCGCAAACCGGCAGTGCGACAGATGCGCCGCGCCGCCGTTCGCGCCGT 651149
Qy 2788 ---cgccgttccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacg 2844
Db 651148 TCGCGCCGTTCCCTATTATCCGTTACACCGCCAACTTCGGTAGAATCCCGTTTCAACACG 651089
Qy 2845 ctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttc 2904
Db 651088 CTGACGGTAAACGGCAAATTGAACGGTCAGGGAACATTCCGCTTTATGTCGGAACTCTTC 651029
Qy 2905 ggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcg 2964
Db 651028 GGCTACCGCAGCGACAAATTGAAGCTGGCGGAAAGTTCCGAAGGCACTTACACCTTGGCG 650969
Qy 2965 gtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaa 3024
Db 650968 GTCAACAATACCGGCAACGAACCTGCAAGCCTCGAACAATTGACGGTAGTGGAAGGAAAA 650909
Qy 3025 gacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcc 3084
Db 650908 GACAACAAACCGCTGTCCGAAAACCTTAATTTCACCCTGCAAAACGAACACGTCGATGCC 650849
Qy 3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
Db 650848 GGCGCGTGGCGTTACCAACTCATCCGCAAAGACGGCGAGTTCCGCCTGCATAATCCGGTC 650789
Qy 3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204
Db 650788 AAAGAACAAGAGCTTTCCGACAAACTCGGCAAGGCAGAAGCCAAAAAACAGGCGGAAAAA 650729
Qy 3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
Db 650728 GACAACGCGCAAAGCCTTGACGCGCTGATTGCGGCCGGGCGCGATGCCGTCGAAAAGACA 650669
Qy 3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
Db 650668 GAAAGCGTTGCCGAACCGGCCCGGCAGGCAGGCGGGGAAAATGTCGGCATTATGCAGGCG 650609
Qy 3325 gaggaagagaaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
Db 650608 GAGGAAGAGAAAAAACGGGTGCAGGCGGATAAAGACACCGCCTTGGCGAAACAGCGCGAA 650549
Qy 3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
Db 650548 GCGGAAACCCGGCCGGCTACCACCGCCTTCCCCCGCGCCCGCCGCGCCCGCCGGGATTTG 650489
Qy 3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
Db 650488 CCGCAACTGCAACCCCAACCG------CAGCCCCAACCGCAGCGCGACCTGATCAGCCGT 650435
Qy 3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
Db 650434 TATGCCAATAGCGGTTTGAGTGAATTTTCCGCCACGCTCAACAGCGTTTTCGCCGTACAG 650375
Qy 3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
Db 650374 GACGAATTAGACCGCGTATTTGCCGAAGACCGCCGCAACGCCGTTTGGACAAGCGGCATC 650315
Qy 3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
Db 650314 CGGGACACCAAACACTACCGTTCGCAAGATTTCCGCGCCTACCGCCAACAAACCGACCTG 650255
Qy 3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
Db 650254 CGCCAAATCGGTATGCAGAAAAACCTCGGCAGCGGGCGCGTCGGCATCCTGTTTTCGCAC 650195
Qy 3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
Db 650194 AACCGGACCGAAAACACCTTCGACGACGGCATCGGCAACTCGGCACGGCTTGCCCACGGC 650135
Qy 3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
Db 650134 GCCGTTTTCGGGCAATACGGCATCGACAGGTTCTACATCGGCATCAGCGCGGGCGCGGGT 650075
Qy 3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
Db 650074 TTTAGCAGCGGCAGCCTTTCAGACGGCATCGGAGGCAAAATCCGCCGCCGCGTGCTGCAT 650015
Qy 3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
Db 650014 TACGGCATTCAGGCACGATACCGCGCCGGTTTCGGCGGATTCGGCATCGAACCGCACATC 649955
Qy 3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
Db 649954 GGCGCAACGCGCTATTTCGTCCAAAAAGCGGATTACCGCTACGAAAACGTCAATATCGCC 649895
Qy 4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
Db 649894 ACCCCCGGCCTTGCATTCAACCGCTACCGCGCGGGCATTAAGGCAGATTATTCATTCAAA 649835
Qy 4105 ccggcgcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcg 4164
Db 649834 CCGGCGCAACACATTTCCATCACGCCTTATTTGAGCCTGTCCTATACCGATGCCGCTTCG 649775
Qy 4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
Db 649774 GGCAAAGTCCGAACACGCGTCAATACCGCCGTATTGGCTCAGGATTTCGGCAAAACCCGC 649715

```
Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db 649714 AGTGCGGAATGGGGCGTAAACGCCGAAATCAAAGGTTTCACGCTGTCCCTCCACGCTGCC 649655

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db 649654 GCCGCCAAAGGCCCGCAACTGGAAGCGCAACACAGCGCGGGCATCAAATTAGGCTACCGC 649595

Qy   4345 tggtaa 4350
          ||||||
Db 649594 TGGTAA 649589
```

RESULT 9
AAZ38918
ID AAZ38918 standard; DNA; 4374 BP.
XX
AC AAZ38918;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain H44/76 BASB006 nucleotide sequence.
XX
KW Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KW vaccine; antibiotic; upper respiratory tract infection; meningitis;
KW invasive bacterial disease; bacteraemia; screening; antibacterial; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1..4374
FT /*tag= a
FT /product= "BASB006"
XX
PN WO9955873-A2.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP02766.
XX
PR 24-APR-1998; 98GB-0008866.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Thonnard J;
XX
DR WPI; 2000-052810/04.
DR P-PSDB; AAY56622.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitis used
PT to prepare vaccines against bacterial infections -
XX
PS Claim 11; Page 91-93; 103pp; English.
XX
CC The present sequence encodes BASB006 isolated from Neisseria meningitidis
CC strain H44/76. BASB006 polynucleotides and polypeptides may be
CC employed as research reagents and material for the discovery of
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of disease, staging of disease, or
CC determining response of an infectious organism to drugs. The
CC polynucleotides may be used as a source for hybridisation probes, and for
CC screening of genetic mutations, serotype, organism or strain
CC identification, identification of mutation in BASB013 sequences, and as
CC components of arrays which are useful for diagnostic and prognostic
CC purposes. The polypeptides can be used to produce antibodies. The
CC polypeptides can also be used in vaccine formulations, and to identify
CC agonists and antagonists. The polypeptides, antibodies, agonists and
CC antagonists (which are bacteristatic) are used for the treatment and
CC prevention of diseases such as upper respiratory tract infection,
CC invasive bacterial diseases such as bacteraemia and meningitis, and for
CC the development and screening of antibacterial drugs. They are also used
CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC prevent tissue damage and/or block the normal progression of pathogenesis
CC in infections initiated other than by the implantation of in-dwelling
CC devices or by other surgical techniques.
XX
SQ Sequence 4374 BP; 1269 A; 1158 C; 1066 G; 881 T; 0 other;

Query Match 80.4%; Score 3498.4; DB 21; Length 4374;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 3870; Conservative 0; Mismatches 468; Indels 48; Gaps 8;

```
Qy      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60

Qy     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagcc 120

Qy    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180

Qy    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 240

Qy    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300

Qy    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360

Qy    361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
          |||||||||||||||||||||||| |||||||||| || |||||| |||| |||||
Db    361 aacgttgattttggtgcggaaggaagaaatcccgatcaacatcgttttacttataaaatt 420

Qy    421 gtgaaaagaaataattataagcctgacaattca---caccctacaacggcgattancat 477
          |||||| | |||||||||| | | | | | || ||||| |||||||| |||
Db    421 gtgaaacggaataattataaagcagggactaaaggccatccttatggtggcgattatcat 480

Qy    478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
          |||||||||| |||||||||||||||||||||||||||| |||||||| ||| | |||
Db    481 atgccgcgtttacataaatttgtcacagatgcagaacctgttgaaatgaccagttatatg 540

Qy    538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
             |||   | |||  |||| || | || || || || ||||| || || ||  ||||
Db    541 gatgggcggaaatatatcgatcaaaataattaccctgaccgtgttcgtattggggcaggc 600

Qy    598 caccactattggcgttatgatgatga------caaacacggcgatttatcctactccggc 651
             || |||||||| | |||||| ||      |||  || ||||    || ||
Db    601 aggcaatattggcgatctgatgaagatgagcccaataaccgcgaaagttcatatcatatt 660

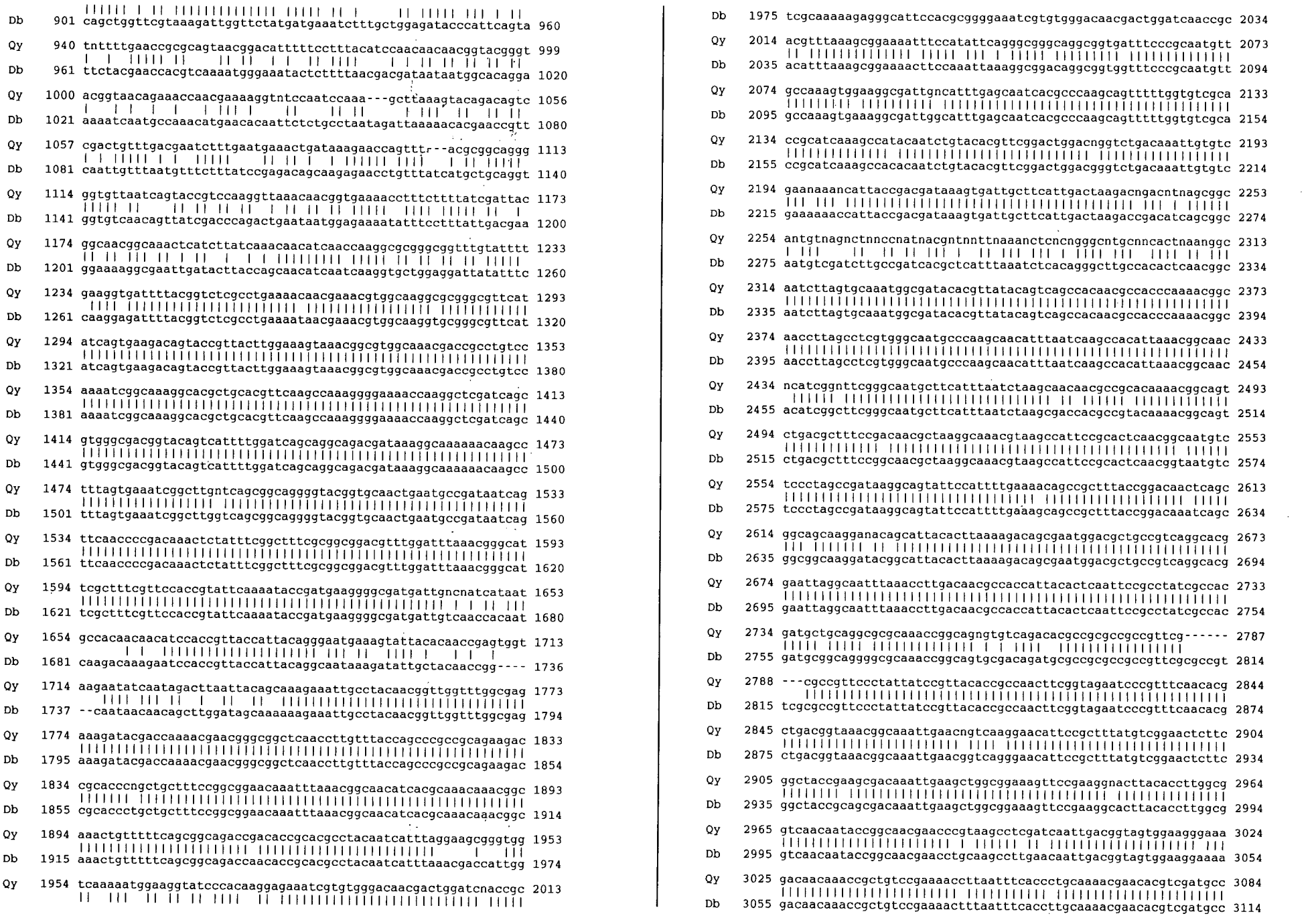
Qy    652 gca----------tggttaattggcggcaatacacatatgcagggttggggaaataat 699
          |||            ||| |  |||| |||||||   |   ||  | |   |  |  |
Db    661 gcaagtgcgtattcttggctcgttggtggcaatacctttgcacaaaatggatcaggtggt 720

Qy    700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
          |||  | | |  ||  |  | ||       |    | |   |||||   | | || |
Db    721 ggcacagtcaacttaggtagtgaaaaaattaaacatagcccatatggtttttttaccaaca 780

Qy    760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
          | |||  |    |||||||| || || ||||||||||| |||||      |||  | ||
Db    781 ggaggctcatttggcgacagtggctcaccaatgtttatctatgatgcccaaaagcaaaag 840

Qy    820 tggctgctcaacggagttttacaaaccggctaccttattccggcagggaaaacggtttc 879
          ||| |  | || || || || ||||| ||| |||| |||   || |     || || |||
Db    841 tggttaattaatggggtattgcaaacgggcaacccctatataggaaaaagcaatggcttc 900

Qy    880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
```



```
Qy   3085 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3144
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3115 ggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtc 3174

Qy   3145 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3175 aaagaacaagagctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaa 3234

Qy   3205 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagaca 3264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3235 gacaacgcgcaaagccttgacgcgctgattgcggccgggcgcgatgccgtcgaaaagaca 3294

Qy   3265 gaaagcgttgccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcg 3324
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   3295 gaaagcgttgccgaaccggcccggcaggcaggcggggaaaatgtcggcattatgcaggcg 3354

Qy   3325 gaggaagagaaaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaa 3384
          |||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||||||
Db   3355 gaggaagagaaaaaacgggtgcaggcggataaagacaccgccttggcgaaacagcgcgaa 3414

Qy   3385 gcggaaacccggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttg 3444
          ||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db   3415 gcggaaacccggccggctaccaccgccttcccccgcgcccgccgcgcccgccgggatttg 3474

Qy   3445 ccgcaaccgcagccccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgt 3504
          ||||||| ||| ||||||||| || |||||||||||||||||||||||| ||||||
Db   3475 ccgcaactgcaaccccaaccg------cagccccaaccgcagcgcgacctgatcagccgt 3528

Qy   3505 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3529 tatgccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacag 3588

Qy   3565 gacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatc 3624
          |||||||| |||||||| |||||||||| |||||||||| |||||||||| |||||
Db   3589 gacgaattagaccgcgtatttgccgaagaacgccgcaacgccgtttggacaagcggcatc 3648

Qy   3625 cggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3684
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3649 cgggacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctg 3708

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3709 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3768

Qy   3745 aaccggaccgaaaacanctttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   3769 aaccggaccgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggc 3828

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
          |||||||||||||||||||||||| ||||||| |||||||||||||| ||||||||||||
Db   3829 gccgttttcgggcaatacggcatcgacaggttctacatcggcatcagcgcgggcgcgggt 3888

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
          |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||||||
Db   3889 tttagcagcggcagcctttcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3948

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   3949 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatc 4008

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4068

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
          |||||||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||||
Db   4069 acccccggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaa 4128

Qy   4105 ccggcgcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcg 4164
          ||||||||||||||| |||||||| |||||||| ||||||||||||||||||||||||||||||
Db   4129 ccggcgcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcg 4188

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   4189 ggcaaagtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgc 4248

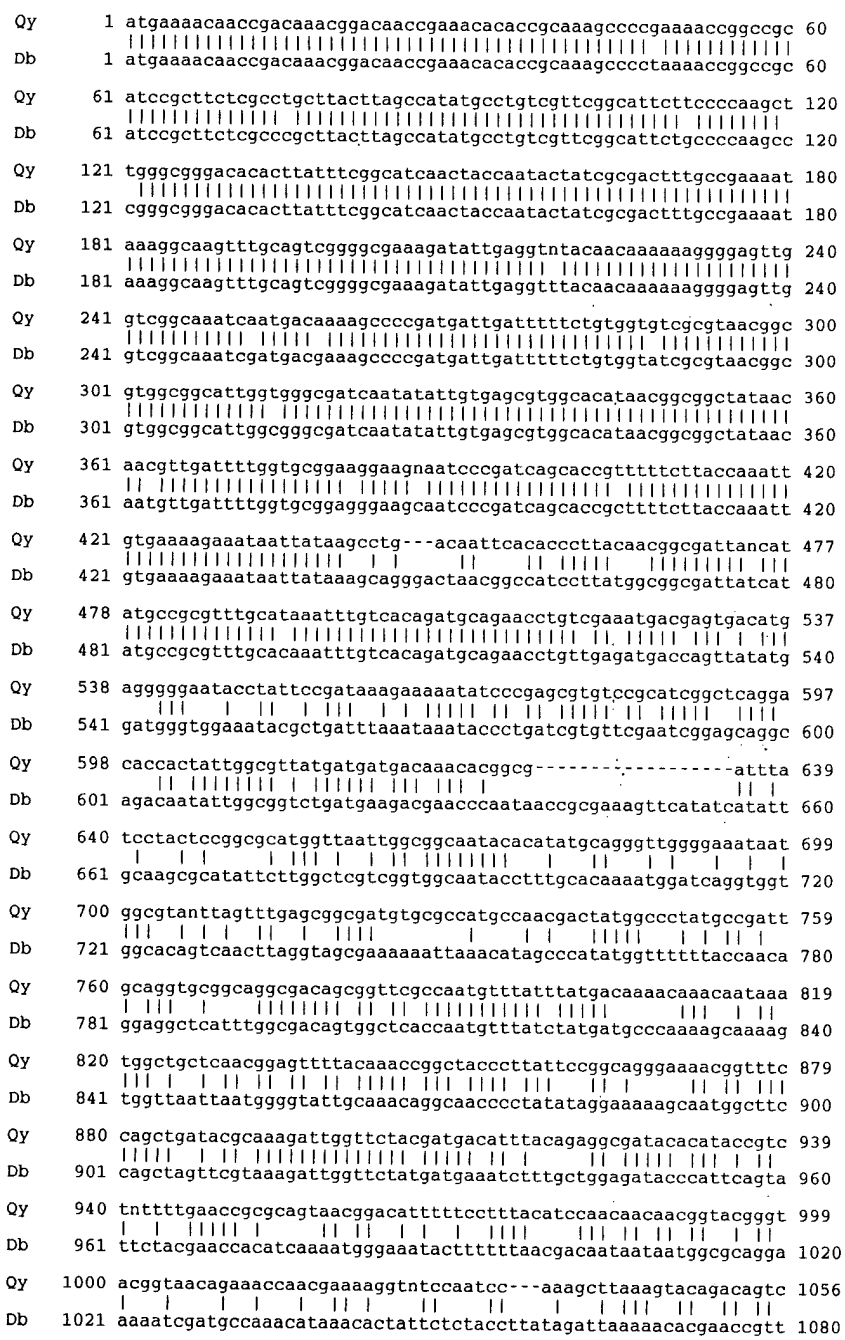
Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
          |||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db   4249 agtgcggaatggggcgtaaacgccgaaatcaaaggcttcacgctgtccctccacgctgcc 4308

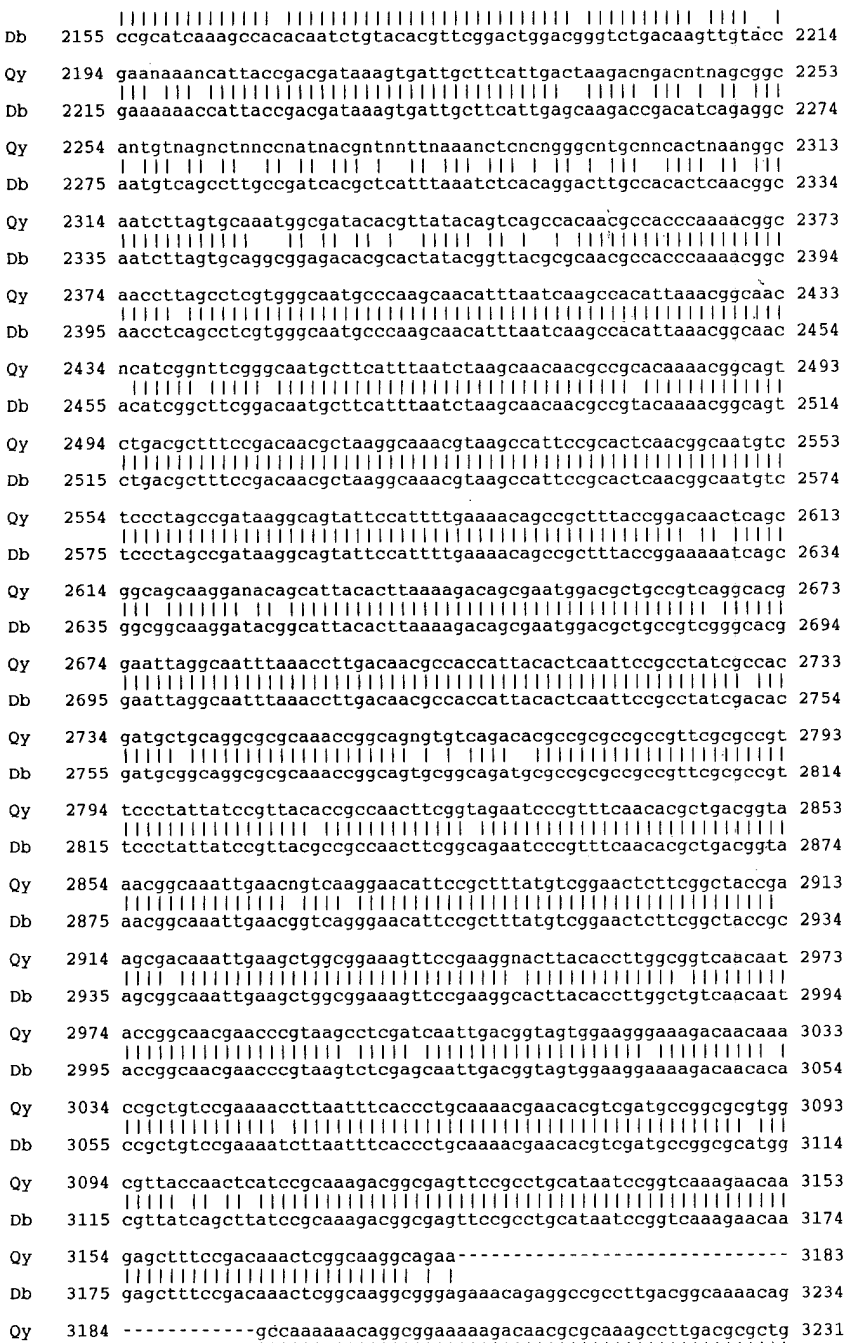
Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 gccgccaaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4368

Qy   4345 tggtaa 4350
          ||||||
Db   4369 tggtaa 4374
```

RESULT 10
AAZ12253
ID AAZ12253 standard; DNA; 4407 BP.
XX
AC AAZ12253;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria gonorrhoeae complete ORF1 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38825.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 370-371; 524pp; English.
XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 4407 BP; 1273 A; 1169 C; 1094 G; 871 T; 0 other;

Query Match 77.8%; Score 3385.8; DB 20; Length 4407;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 3814; Conservative 0; Mismatches 524; Indels 81; Gaps 7;


Qy 1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
Db 1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagcccctaaaaccggccgc 60
Qy 61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
Db 61 atccgcttctcgcccgcttacttagccatatgcctgtcgttcggcattctgccccaagcc 120
Qy 121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
Db 121 cgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
Qy 181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
Db 181 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 240
Qy 241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
Db 241 gtcggcaaatcgatgacgaaagccccgatgattgatttttctgtggtatcgcgtaacggc 300
Qy 301 gtggcggcattggtgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
Db 301 gtggcggcattggcgggcgatcaatatattgtgagcgtggcacataacggcggctataac 360
Qy 361 aacgttgattttggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaatt 420
Db 361 aatgttgattttggtgcggagggaagcaatcccgatcagcaccgcttttcttaccaaatt 420
Qy 421 gtgaaaagaaataattataagcctg---acaattcacaccttacaacggcgattancat 477
Db 421 gtgaaaagaaataattataaagcagggactaacggccatccttatggcggcgattatcat 480
Qy 478 atgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatg 537
Db 481 atgccgcgtttgcacaaatttgtcacagatgcagaacctgttgagatgaccagttatatg 540
Qy 538 agggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcagga 597
Db 541 gatgggtggaaatacgctgatttaaataaataccctgatcgtgttcgaatcggagcaggc 600
Qy 598 caccactattggcgttatgatgatgacaaacacggcg--------atta 639
Db 601 agacaatattggcggtctgatgaagacgaacccaataaccgcgaaagttcatatcatatt 660
Qy 640 tcctactccggcgcatggttaattggcggcaatacacatatgcagggttggggaaataat 699
Db 661 gcaagcgcatattcttggctcgtcggtggcaatacctttgcacaaaatggatcaggtggt 720
Qy 700 ggcgtanttagtttgagcggcgatgtgcgccatgccaacgactatggccctatgccgatt 759
Db 721 ggcacagtcaacttaggtagcgaaaaaattaaacatagcccatatggtttttaccaaca 780
Qy 760 gcaggtgcggcaggcgacagcggttcgccaatgtttatttatgacaaaacaaacaataaa 819
Db 781 ggaggctcatttggcgacagtggctcaccaatgtttatctatgatgcccaaaagcaaaag 840
Qy 820 tggctgctcaacggagttttacaaaccggctacccttattccggcagggaaaacggtttc 879
Db 841 tggttaattaatggggtattgcaaacaggcaacccctatataggaaaaagcaatggcttc 900
Qy 880 cagctgatacgcaaagattggttctacgatgacatttacagaggcgatacacataccgtc 939
Db 901 cagctagttcgtaaagattggttctatgatgaaatctttgctggagatacccattcagta 960
Qy 940 tnttttgaaccgcgcagtaacggacatttttccttttacatccaacaacaacggtacgggt 999
Db 961 ttctacgaaccacatcaaaatgggaaatacttttttaacgacaataataatggcgcagga 1020
Qy 1000 acggtaacagaaaccaacgaaaaggtntccaatcc---aaagcttaaagtacagacagtc 1056
Db 1021 aaaatcgatgccaaacataaacactattctctaccttatagattaaaaacacgaaccgtt 1080
Qy 1057 cgactgtttgacgaatctttgaatgaaactgataaagaaccagttt---acgcggcaggg 1113
Db 1081 caattgtttaatgtttctttatccgagacagcaagagaacctgtttatcatgctgcaggt 1140
Qy 1114 ggtgttaatcagtaccgtccaaggttaaacaacggtgaaaacctttctttttatcgattac 1173
Db 1141 ggggtcaacagttatcgacccagactgaataatggagaaaatatttcctttattgacaaa 1200
Qy 1174 ggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttt 1233
Db 1201 ggaaaaggtgaattgatacttaccagcaacatcaaccaaggcgcgggcggtttgtatttt 1260
Qy 1234 gaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcat 1293
Db 1261 gagggtaattttacggtctcgcctaaaaacaacgaaacgtggcaaggcgcgggcgttcat 1320
Qy 1294 atcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1353
Db 1321 atcagtgatggcagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtcc 1380
Qy 1354 aaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagc 1413
Db 1381 aaaatcggcaaaggcacgctgctggttcaagccaaaggggaaaaccaaggctcggtcagc 1440
Qy 1414 gtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagcc 1473
Db 1441 gtgggcgacggtaaagtcatcttagatcagcaggcggacgatcaaggcaaaaaacaagcc 1500
Qy 1474 tttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcag 1533
Db 1501 tttagtgaaatcggcttggtcagcggcaggggacggtgcaactgaatgccgataatcag 1560
Qy 1534 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcat 1593
Db 1561 ttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttgaacgggcat 1620
Qy 1594 tcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcataat 1653
Db 1621 tcgctttcgttccaccgcattcaaaataccgatgaaggggcgatgattgtcaaccacaat 1680
Qy 1654 gccacaacaacatccaccgttaccattacagggaatgaaagtattacacaaccgagtggt 1713
Db 1681 caagacaaagaatccaccgttaccattacaggcaataaagatattactacaaccggc--- 1737
Qy 1714 aagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtttggcgag 1773
Db 1738 ---aataacaacaacttggatagcaaaaaagaaattgcctacaacggttggtttggcgag 1794
Qy 1774 aaagatacgaccaaaacgaacgggcggctcaaccttgtttaccagcccgccgcagaagac 1833
Db 1795 aaagatgcaaccaaaacgaacgggcggctcaatctgaattaccaaccggaagaagcggat 1854
Qy 1834 cgcacccngctgctttccggcggaacaaatttaaacggcaacatcacgcaaacaaacggc 1893
Db 1855 cgcactttactgctttccggcggaacaaatttaaacggcaatatcacgcaaacaaacggc 1914
Qy 1894 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1953
Db 1915 aaactgtttttcagcggcagaccgacaccgcacgcctacaatcatttaggaagcgggtgg 1974
Qy 1954 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgactggatcnaccgc 2013
Db 1975 tcaaaaatggaaggtatcccacaaggagaaatcgtgtgggacaacgattggatcgaccgc 2034
Qy 2014 acgtttaaagcggaaaatttccatattcagggcgggcaggcggtgatttcccgcaatgtt 2073
Db 2035 acatttaaagcggaaaacttccatattcagggcggacaagcggtggtttcccgcaatgtt 2094
Qy 2074 gccaaagtggaaggcgattgncatttgagcaatcacgcccaagcagttttggtgtcgca 2133
Db 2095 gccaaagtggaaggcgattggcatttaagcaatcacgcccaagcagttttcggtgtcgca 2154
Qy 2134 ccgcatcaaagccatacaatctgtacacgttcggactggacnggtctgacaaattgtgtc 2193


Db 2155 ccgcatcaaagccacacaatctgtacacgttcggactggacgggtctgacaagttgtacc 2214
Qy 2194 gaanaaancattaccgacgataaagtgattgcttcattgactaagacngacntnagcggc 2253
Db 2215 gaaaaaaccattaccgacgataaagtgattgcttcattgagcaagaccgacatcagaggc 2274
Qy 2254 antgtnagnctnnccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggc 2313
Db 2275 aatgtcagccttgccgatcacgctcatttaaatctcacaggacttgccacactcaacggc 2334
Qy 2314 aatcttagtgcaaatggcgatacacgttatacagtcagccacaacgccacccaaaacggc 2373
Db 2335 aatcttagtgcaggcggagacacgcactatacggttacgcgcaacgccacccaaaacggc 2394
Qy 2374 aaccttagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2433
Db 2395 aacctcagcctcgtgggcaatgcccaagcaacatttaatcaagccacattaaacggcaac 2454
Qy 2434 ncatcggnttcgggcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagt 2493
Db 2455 acatcggcttcggacaatgcttcatttaatctaagcaacaacgccgtacaaaacggcagt 2514
Qy 2494 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2553
Db 2515 ctgacgctttccgacaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtc 2574
Qy 2554 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggacaactcagc 2613
Db 2575 tccctagccgataaggcagtattccattttgaaaacagccgctttaccggaaaaatcagc 2634
Qy 2614 ggcagcaaggganacagcattacacttaaaagacagcgaatggacgctgccgtcaggcacg 2673
Db 2635 ggcggcaaggatacggcattacacttaaaagacagcgaatggacgctgccgtcgggcacg 2694
Qy 2674 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgccac 2733
Db 2695 gaattaggcaatttaaaccttgacaacgccaccattacactcaattccgcctatcgacac 2754
Qy 2734 gatgctgcaggcgcgcaaaccggcagngtgtcagacacgccgcgccgccgttcgcgccgt 2793
Db 2755 gatgcggcaggcgcgcaaaccggcagtgcggcagatgcgccgcgccgccgttcgcgccgt 2814
Qy 2794 tccctattatccgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggta 2853
Db 2815 tccctattatccgttacgccgccaacttcggcagaatcccgtttcaacacgctgacggta 2874
Qy 2854 aacggcaaattgaacngtcaaggaacattccgctttatgtcggaactcttcggctaccga 2913
Db 2875 aacggcaaattgaacggtcagggaacattccgctttatgtcggaactcttcggctaccgc 2934
Qy 2914 agcgacaaattgaagctggcggaaagttccgaaggnacttacaccttggcggtcaacaat 2973
Db 2935 agcggcaaattgaagctggcggaaagttccgaaggcacttacaccttggctgtcaacaat 2994
Qy 2974 accggcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaagacaacaaa 3033
Db 2995 accggcaacgaacccgtaagtctcgagcaattgacggtagtggaaggaaaagacaacaca 3054
Qy 3034 ccgctgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgccggcgcgtgg 3093
Db 3055 ccgctgtccgaaaatcttaatttcaccctgcaaaacgaacacgtcgatgccggcgcatgg 3114
Qy 3094 cgttaccaactcatccgcaaagacggcgagttccgcctgcataatccggtcaaagaacaa 3153
Db 3115 cgttatcagcttatccgcaaagacggcgagttccgcctgcataatccggtcaaagaacaa 3174
Qy 3154 gagctttccgacaaactcggcaaggcagaa---------- 3183
Db 3175 gagctttccgacaaactcggcaaggcgggagaaacagaggccgccttgacggcaaaacag 3234
Qy 3184 ----------gccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231
Db 3235 gcacaacttgccgccaaacaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3294
Qy 3232 attgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcng 3291
Db 3295 attgcggccgggcgcaatgccaccgaaaaggcagaaagtgttgccgaaccggcccggcag 3354
Qy 3292 gcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3351
Db 3355 gcaggcggggaaaatgccggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3414
Qy 3352 gataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgcc 3411
Db 3415 gataaagacaccgccttggcgaaacagcgcgaagcggaaacccggccggctaccaccgcc 3474
Qy 3412 ttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagcccccaaccgcaacct 3471
Db 3475 ttcccccgcgcccgccgcgcccgccgggatttgccgcaaccgcagcccccaaccg------ 3528
Qy 3472 caaccccaaccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattt 3531
Db 3529 caaccccaaccgcagcgcgacctgatcagccgttatgccaatagcggtttgagtgaattt 3588
Qy 3532 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3591
Db 3589 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3648
Qy 3592 gaccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaa 3651
Db 3649 gaccgccgcaacgccgtttggacaagcggcatccgggacaccaaacactaccgttcgcaa 3708
Qy 3652 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3711
Db 3709 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3768
Qy 3712 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgac 3771
Db 3769 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccggaaacaccttcgacgac 3828
Qy 3772 ggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggc 3831
Db 3829 ggcatcggcaactcggcacggcttgcccacggtgccgttttcgggcaatacggcatcggc 3888
Qy 3832 aggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggc 3891
Db 3889 aggttcgacatcggcatcagcgcgggcgcgggttttagtagcggcagcctttcagacggc 3948
Qy 3892 atcggaggcaaaatccgccgccgcgtgctgcattacggcattcaggcacgataccgcgcc 3951
Db 3949 atcagaggcaaaatccgccgccgcgtgctgcattacggcattcaggcaagataccgcgca 4008
Qy 3952 ggtttcggcggattcggcatcgaaccgtacatcggcgcaacgcgctatttcgtccaaaaa 4011
Db 4009 ggtttcggcggattcggcatcgaaccgcacatcggcgcaacgcgctatttcgtccaaaaa 4068
Qy 4012 gcggattaccgctacgaaaacgtcaatatcgccaccccggtcttgcgttcaaccgntac 4071
Db 4069 gcggattaccgatacgaaaacgtcaatatcgccaccccgggccttgcattcaaccgctac 4128
Qy 4072 cgngcgggcattaaggcagattattcattcaaaccggcgcaacacatntccatcacncct 4131
Db 4129 cgcgcgggcattaaggcagattattcattcaaaccggcgcaacacatttccatcacgcct 4188
Qy 4132 tatttnagcctgtcctataccgatgccgcttcgggcaaagtccgaacacgcgtcaatacc 4191
Db 4189 tatttgagcctgtcctataccgatgccgcttccggcaaagtccgaacgcgcgtcaatacc 4248
Qy 4192 gcngtattggctcaggatttcggcaaaacccgcagtgcggaatggggcgtaaacgccgaa 4251
Db 4249 gccgtattggcgcaggatttcggcaaaacccgcagtgcggaatggggcgtaaacgccgaa 4308
Qy 4252 atcaaaggtttcacgctgtccntccacgctgccgccgccaaaggnccgcaactggaagcg 4311
Db 4309 atcaaaggtttcacgctgtccctccacgctgccgccgccaaggggccgcaattggaagcg 4368

```
Qy   4312 caacacagcgcgggcatcaaattaggctaccgctggtaa 4350
          || ||||||||||||||||||||||||||||||||||||
Db   4369 cagcacagcgcgggcatcaaattaggctaccgctggtaa 4407


RESULT  11
AAT17215
ID   AAT17215 standard; DNA; 4319 BP.
XX
AC   AAT17215;
XX
DT   01-JUN-1996  (first entry)
XX
DE   Adhesion and penetration protein gene.
XX
KW   Adhesion and penetration protein; hap gene; promoter; terminator;
KW   inverted repeat; stem-loop; protease; outer membrane protein;
KW   beta-domain; secretion; recombinant vaccine; monoclonal antibody;
KW   diagnostic; immunoassay; ds.
XX
OS   Haemophilus influenzae.
XX
FH   Key             Location/Qualifiers
FT   -35_signal      5..10
FT                   /*tag=  a
FT   -10_signal      26..31
FT                   /*tag=  b
FT   sig_peptide     60..134
FT                   /*tag=  c
FT   CDS             60..4244
FT                   /*tag=  d
FT                   /product= Adhesion and penetration preprotein
FT   mat_peptide     135..4241
FT                   /*tag=  e
FT                   /product= Adhesion and penetration protein
FT   mat_peptide     135..2936
FT                   /*tag=  f
FT                   /product= Secreted 110-kDa protease fragment
FT   mat_peptide     2937..4241
FT                   /*tag=  g
FT                   /product= 45-kDa outer membrane protein fragment
FT   terminator      4284..4302
FT                   /*tag=  h
FT   repeat_region   4284..4302
FT                   /*tag=  i
FT                   /rpt_type= INVERTED
FT   stem_loop       4284..4302
FT                   /*tag=  j
XX
PN   WO9605858-A1.
XX
PD   29-FEB-1996.
XX
PF   16-AUG-1995;   95WO-US10661.
XX
PR   25-AUG-1994;   94US-0296791.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Falkow S,  St Geme JW;
XX
DR   WPI; 1996-151147/15.
DR   P-PSDB; AAR92768.
XX
PT   Haemophilus adhesion and penetration protein and corresponding DNA
PT   - used to produce vaccines against H. influenzae infection
XX
PS   Claim 5; Fig 6; 105pp; English.
XX
CC   The sequence encodes a Haemophilus influenzae adhesion and
CC   penetration protein.  The sequence (hap gene) includes putative
CC   -10 and -30 sequences and a putative rho-independent terminator 3'
CC   to the hap stop codon.  The terminator contains interrupted
CC   inverted repeats, with the potential for forming a hairpin
CC   structure containing a loop of 3 bases and a stem of 8 bases,
CC   followed by a stretch rich in T residues.  The gene product is
CC   first synthesised as a preprotein, which is transported to the
CC   periplasm, followed by insertion of the C-terminal beta-domain into
CC   the outer membrane, possibly forming a pore, and export of the
CC   N-terminal fragment through the outer membrane, followed by
CC   autoproteolytic cleavage and secretion of the mature protease, leaving
CC   an outer membrane protein fragment.  The gene may be inserted in a
CC   vector and expressed in recombinant host cells, for use as a
CC   recombinant vaccine.  The gene product may also be used in
CC   diagnostic monoclonal antibody production.
XX
SQ   Sequence 4319 BP; 1497 A; 776 C; 891 G; 1155 T; 0 other;


  Query Match             31.0%;  Score 1348;  DB 17;  Length 4319;
  Best Local Similarity   59.0%;  Pred. No. 0;
  Matches 2551;  Conservative    0;  Mismatches 1565;  Indels  210;  Gaps    9;

Qy     73 cctgcttacttagccatatgcctgtcgttcggcattcttccccaagcttgggcgggacac 132
          | |  ||  ||| ||   ||| | || || || ||  |  | ||||| ||||| || |||
Db     81 cttaatttttaaccgcttgcatttcattagggatagtatcgcaagcgtgggctggtcac 140

Qy    133 acttatttcggcatcaactaccaatactatcgcgactttgccgaaaataaaggcaagttt 192
          |||||||| || ||  | |||||||| ||||| || |||||||| |||||||| |||||
Db    141 acttattttgggattgattaccaatattatcgtgattttgccgagaataaagggaagttc 200

Qy    193 gcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttggtcggcaaatca 252
           |||| |||||  || ||||| |||| || |||||| ||||| | || || |||| ||||
Db    201 acagttggggctcaaaatattaaggtttataacaaacaagggcaattagttggcacatca 260

Qy    253 atgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggcgtggcggcattg 312
          ||||||||||||||||||||||||||||||||| ||||| ||||||||||||||| || |||
Db    261 atgacaaaagccccgatgattgatttttctgtagtgtcacgtaacggcgtggcagccttg 320

Qy    313 gtgggcgatcaatatattgtgagcgtggcacataacggcggctataacaacgttgatttt 372
          || |  |||||||||||||||||||||||||||||  || ||||  | |||||||||
Db    321 gttgaaaatcaatatattgtgagcgtggcacataacgtaggatatacagatgttgatttt 380

Qy    373 ggtgcggaaggaagnaatcccgatcagcaccgtttttcttaccaaattgtgaaaagaaat 432
          ||||| || |||| || |||||||| || |||||| |||| | ||||| ||| |||||
Db    381 ggtgcagagggaaacaaccccgatcaacatcgttttacttataagattgtaaaacgaaat 440

Qy    433 aattataagcctgacaattcacacccttacaacggcgattancatatgccgcgtttgcat 492
          || || ||    || |||| ||| |||||  | | |||||| |||| || || || |||
Db    441 aactacaaaaaagatatttacatccttatgaggacgattaccataatccacgattacat 500

Qy    493 aaatttgtcacagatgcagaacctgtcgaaatgacgagtgacatgaggggaatacctat 552
          ||||| || ||||| || |  ||  | || |||||     | ||||  || | ||| |||
Db    501 aaattcgttacagaagcggctccaattgatatgacttcgaatatgaatggcagtacttat 560

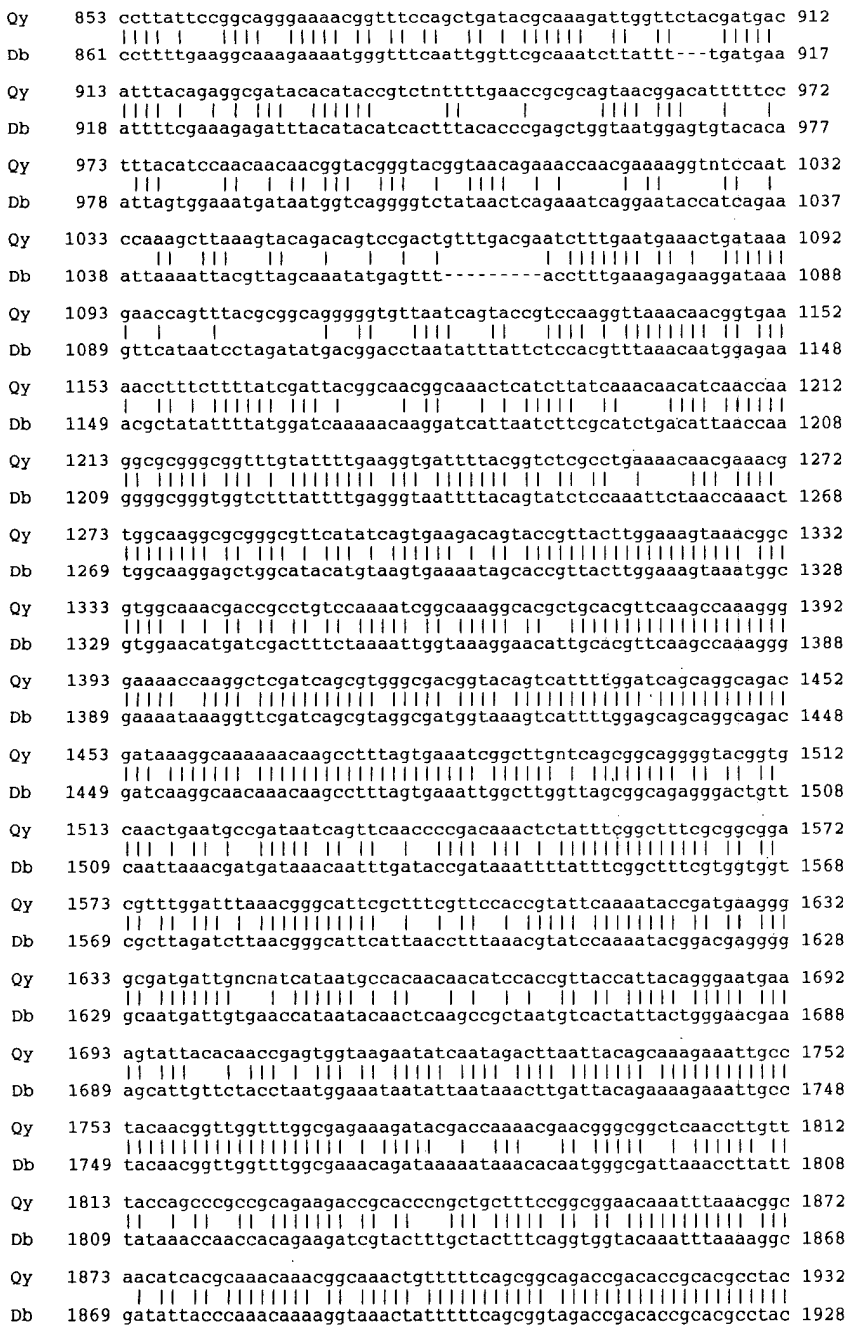
Qy    553 tccgataaagaaaaatatcccgagcgtgtccgcatcggctcaggacaccactattggcgt 612
          || |||| |  |||||||||| || ||||| || |||||||| |||| || | ||||||
Db    561 tcagatagaacaaaatatccagaacgtgttcgtatcggctctggacggcagttttggcga 620

Qy    613 tatgatgatgacaaacacggcgatttatcctactccggcgcatggttaattggcggcaat 672
           ||||| | ||||||  || | |  |  ||    |       |  | |  |  ||||||
Db    621 aatgatcaagacaaaggcgaccaagttgccggtgcatatcattatctgacagctggcaat 680

Qy    673 acacatatgcagggttggggaaataatggcgtanttagtttgagcggcgatgtgcgccat 732
          ||||| |  ||| || | | |  ||||||         |||| | |||||||| ||  |
Db    681 acacacaatcagcgtggagcaggtaatggatattcgtatttgggaggcgatgttcgtaaa 740

Qy    733 gccaacgactatggccctatgccgattgcaggtgcggcaggcgacagcggttcgccaatg 792
          ||    || ||||| ||  | ||||||||||  |    || ||||| ||||| || |||
Db    741 gcgggagaatatggtccattaccgattgcaggctcaaaggggacagtggttctccgatg 800

Qy    793 tttatttatgacaaaacaaacaataaatggctgctcaacggagttttacaaaccggctac 852
          ||||||||||       |||  | ||||||  |  | || ||  | ||||     ||| ||
Db    801 tttatttatgatgctgaaaaacaaaaatggttaattaatgggatattacgggaaggcaac 860
```


Qy 853 ccttattccggcagggaaaacggtttccagctgatacgcaaagattggttctacgatgac 912
Db 861 ccttttgaaggcaaagaaaatgggtttcaattggttcgcaaatcttattt---tgatgaa 917
Qy 913 atttacagaggcgatacacataccgtctnttttgaaccgcgcagtaacggacatttttcc 972
Db 918 attttcgaaagagatttacatacatcactttacacccgagctggtaatggagtgtacaca 977
Qy 973 tttacatccaacaacaacggtacgggtacggtaacagaaaccaacgaaaaggtntccaat 1032
Db 978 attagtggaaatgataatggtcaggggtctataactcagaaatcaggaataccatcagaa 1037
Qy 1033 ccaaagcttaaagtacagacagtccgactgtttgacgaatctttgaatgaaactgataaa 1092
Db 1038 attaaaattacgttagcaaatatgagttt---------acctttgaaagagaaggataaa 1088
Qy 1093 gaaccagtttacgcggcaggggtgttaatcagtaccgtccaaggttaaacaacggtgaa 1152
Db 1089 gttcataatcctagatatgacggacctaatatttattctccacgtttaaacaatggagaa 1148
Qy 1153 aacctttcttttatcgattacggcaacggcaaactcatcttatcaaacaacatcaaccaa 1212
Db 1149 acgctatattttatggatcaaaaacaaggatcattaatcttcgcatctgacattaaccaa 1208
Qy 1213 ggcgcgggcggtttgtattttgaaggtgattttacggtctcgcctgaaaacaacgaaacg 1272
Db 1209 ggggcgggtggtctttatttttgagggtaattttacagtatctccaaattctaaccaaact 1268
Qy 1273 tggcaaggcgcgggcgttcatatcagtgaagacagtaccgttacttggaaagtaaacggc 1332
Db 1269 tggcaaggagctggcatacatgtaagtgaaaatagcaccgttacttggaaagtaaatggc 1328
Qy 1333 gtggcaaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgttcaagccaaaggg 1392
Db 1329 gtggaacatgatcgactttctaaaattggtaaaggaacattgcacgttcaagccaaaggg 1388
Qy 1393 gaaaaccaaggctcgatcagcgtgggcgacggtacagtcattttggatcagcaggcagac 1452
Db 1389 gaaaatgaaggttcgatcagcgtaggcgatggtaaagtcattttggagcagcaggcagac 1448
Qy 1453 gataaaggcaaaaaacaagcctttagtgaaatcggcttgntcagcggcaggggtacggtg 1512
Db 1449 gatcaaggcaacaaacaagcctttagtgaaattggcttggttagcggcagagggactgtt 1508
Qy 1513 caactgaatgccgataatcagttcaaccccgacaaactctatttcggctttcgcggcgga 1572
Db 1509 caattaaacgatgataaacaatttgataccgataaatttttatttcggctttcgtggtggt 1568
Qy 1573 cgtttggatttaaacgggcattcgctttcgttccaccgtattcaaaataccgatgaaggg 1632
Db 1569 cgcttagatcttaacgggcattcattaacctttaaacgtatccaaaatacggacgagggg 1628
Qy 1633 gcgatgattgncnatcataatgccacaacaacatccaccgttaccattacagggaatgaa 1692
Db 1629 gcaatgattgtgaaccataatacaactcaagccgctaatgtcactattactgggaacgaa 1688
Qy 1693 agtattacacaaccgagtggtaagaatatcaatagacttaattacagcaaagaaattgcc 1752
Db 1689 agcattgttctacctaatggaaataatattaataaacttgattacagaaaagaaattgcc 1748
Qy 1753 tacaacggttggtttggcgagaaagatacgaccaaaacgaacgggcggctcaaccttgtt 1812
Db 1749 tacaacggttggtttggcgaaacagataaaaataaacacaatgggcgattaaaccttatt 1808
Qy 1813 taccagcccgccgcagaagaccgcacccngctgctttccggcggaacaaatttaaacggc 1872
Db 1809 tataaaccaaccacagaagatcgtactttgctactttcaggtggtacaaatttaaaaggc 1868
Qy 1873 aacatcacgcaaacaaacggcaaactgtttttcagcggcagaccgacaccgcacgcctac 1932
Db 1869 gatattacccaaacaaaaggtaaactatttttcagcggtagaccgacaccgcacgcctac 1928
Qy 1933 aatcatttaggaagcgggtggtcaaaaatggaaggtatcccacaaggagaaatcgtgtgg 1992
Db 1929 aatcatttaaataaacgttggtcagaaatggaaggtataccacaaggcgaaattgtgtgg 1988
Qy 1993 gacaacgactggatcnaccgcacgtttaaagcggaaaatttccatattcagggcgggcag 2052
Db 1989 gatcacgattggatcaaccgtacatttaaagctgaaaacttccaaattaaaggcggaagt 2048
Qy 2053 gcggtgatttcccgcaatgttgccaaagtggaaggcgattgncatttgagcaatcacgcc 2112
Db 2049 gcggtggtttctcgcaatgtttcttcaattgagggaaattggacagtcagcaataatgca 2108
Qy 2113 caagcagtttttggtgtcgcaccgcatcaaagccatacaatctgtacacgttcggactgg 2172
Db 2109 aatgccacatttggtgttgtgccaaatcaacaaaataccatttgcacgcgttcagattgg 2168
Qy 2173 acnggtctgacaaattgtgtcgaanaaancattaccgacgataaagtgattgcttcattg 2232
Db 2169 acaggattaacgacttgtcaaaaagtggatttaaccgatacaaaagttattaattctata 2228
Qy 2233 actaagacngacntnagcggcantgtnagnctnnccnatnacgntnnttnaaanctcncn 2292
Db 2229 ccaaaaacacaaatcaatggctctattaatttaactgataatgcaacggcgaatgttaaa 2288
Qy 2293 gggcntgcnncactnaanggcaatcttagtgcaaatggcgatacacgttatacagtcagc 2352
Db 2289 ggtttagcaaaacttaatggcaatgtcactttaacaaatcacagccaatttacattaagc 2348
Qy 2353 cacaacgccacccaaaacggcaaccttagcctcgtgggcaatgcccaagcaacatttaat 2412
Db 2349 aacaatgccacccaaataggcaatattcg---------- 2377
Qy 2413 caagccacattaaacggcaacncatcggnttcgggcaatgcttcatttaatctaagcaac 2472
Db 2378 ---------- 2377
Qy 2473 aacgccgcacaaaacggcagtctgacgctttccgacaacgctaaggcaaacgtaagccat 2532
Db 2378 ----------actttccgacaattcaactgcaacggtggataat 2411
Qy 2533 tccgcactcaacggcaatgtctccctagccgataaggcagtattccatttgaaaacagc 2592
Db 2412 gcaaacttgaacggtaatgtgcatttaacggattcagctcaattttctttaaaaaacagc 2471
Qy 2593 cgctttaccggacaactcagcggcagcaaggaanacagcattacacttaaaagacagcgaa 2652
Db 2472 catttttcgcaccaaattcagggagacaaaggcacaacagtgacgttggaaaatgcgact 2531
Qy 2653 tggacgctgccgtcaggcacggaattaggcaatttaaaccttgacaacgccaccattaca 2712
Db 2532 tggacaatgcctagcgatactacattgcagaatttaacgctaaataacagtacgatcacg 2591
Qy 2713 ctcaattccgcctatcgccacgatgctgcaggcgcgcaaaccggcagngtgtcagacacg 2772
Db 2592 ttaaattcagcttatt----------cagctagctcaaacaatacg 2627
Qy 2773 ccgcgccgccgttcgcgccgttccctattatccgttacaccgccaacttcggtagaatcc 2832
Db 2628 ccacg---------tcgccgttcattagagacggaaacaacgccaacatcggcagaacat 2678
Qy 2833 cgtttcaacacgctgacggtaaacggcaaattgaacngtcaaggaacattccgctttatg 2892
Db 2679 cgtttcaacacattgacagtaaatggtaaattgagtgggcaaggcacattccaatttact 2738
Qy 2893 tcggaactcttcggctaccgaagcgacaaattgaagctggcggaaagttccgaaggnact 2952
Db 2739 tcatctttatttggctataaaagcgataaattaaaattatccaatgacgctgagggcgat 2798
Qy 2953 tacaccttggcggtcaacaataccggcaacgaacccgtaagcctcgatcaattgacggta 3012
Db 2799 tacatattatctgttcgcaacacaggcaaagaacccgaaacccttgagcaattaactttg 2858
Qy 3013 gtggaagggaaagacaacaaaccgctgtccgaaaaccttaatttcaccctgcaaaacgaa 3072

```
Db   2859 gttgaaagcaaagataatcaaccgttatcagataagctcaaatttactttagaaaatgac 2918

Qy   3073 cacgtcgatgccggcgcgtggcgttaccaactcatccgcaaagacggcgagttccgcctg 3132
Db   2919 cacgttgatgcaggtgcattacgttataaattagtgaagaatgatggcgaattccgcttg 2978

Qy   3133 cataatccggtcaaagaacaagagctttccgacaaactcggcaaggc------------- 3179
Db   2979 cataacccaataaaagagcaggaattgcacaatgatttagtaagagcagagcaagcagaa 3038

Qy   3180 --------agaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231
Db   3039 cgaacattagaagccaaacaagttgaaccgactgctaaaacacaaacaggtgagccaaaa 3098

Qy   3232 attgcggccgggcgcgatgc---cgccgaaaagacagaaagcgttgccgaaccggcccgg 3288
Db   3099 gtgcggtcaagaagagcagcgagagcagcgtttcctgataccctgcctgatcaaagcctg 3158

Qy   3289 cnggcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcag 3348
Db   3159 ttaaacgcattagaagccaaacaagctgaactgactgctgaaacacaaaaaagtaaggca 3218

Qy   3349 gcggataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccacc 3408
Db   3219 aaaacaaaaaaagtgcggtcaaaaagagcagtgttttctgatcccctgcttgatcaaagc 3278

Qy   3409 gccttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaa 3468
Db   3279 ctgttcgcattagaagccgcacttgaggttattgatgccccacagcaatcggaaaaagat 3338

Qy   3469 cctcaaccccaa----------------------ccgcagcgcgacctgatnagccgttat 3507
Db   3339 cgtctagctcaagaagaagcggaaaaacaacgcaaacaaaaagacttgatcagccgttat 3398

Qy   3508 gccaatagcggtttgagtgaattttccgccacgctcaacagcgttttcgccgtacaggac 3567
Db   3399 tcaaatagtgcgttatcagaattatctgcaacagtaaatagtatgctttctgttcaagat 3458

Qy   3568 gaattggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatccgg 3627
Db   3459 gaattagatcgtctttttgtagatcaagcacaatctgccgtgtggacaaatatcgcacag 3518

Qy   3628 nacaccaaacactaccgttcgcaagatttccgcgcctaccgccaac---aaaccgacctg 3684
Db   3519 gataaaagacgctatgattctgatgcgttccgtgcttatcagcagcagaaaacgaacttа 3578

Qy   3685 cgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcac 3744
Db   3579 cgtcaaattggggtgcaaaaagccttagctaatggacgaattggggcagttttctcgcat 3638

Qy   3745 aaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggc 3804
Db   3639 agccgttcagataatacctttgatgaacaggttaaaaatcacgcgacattaacgatgatg 3698

Qy   3805 gccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggt 3864
Db   3699 tcgggttttgcccaatatcaatggggcgatttacaatttggtgtaaacgtgggaacggga 3758

Qy   3865 tttagcagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcat 3924
Db   3759 atcagtgcgagtaaaatggctgaagaacaaagccgaaaaattcatcgaaaagcgataaat 3818

Qy   3925 tacggcattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatc 3984
Db   3819 tatggcgtgaatgcaagttatcagttccgtttagggcaattgggcattcagccttatttt 3878

Qy   3985 ggcgcaacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgcc 4044
Db   3879 ggagttaatcgctatttttattgaacgtgaaaattatcaatctgaggaagtgagagtgaaa 3938

Qy   4045 acccccggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaa 4104
```

```
Db   3939 acgcctagccttgcatttaatcgctataatgctggcattcgagttgattatacatttact 3998

Qy   4105 ccggcgcaacacatntccatcacnccttatttnagcctgtcctataccgatgccgcttcg 4164
Db   3999 ccgacagataatatcagcgttaagccttatttcttcgtcaattatgttgatgtttcaaac 4058

Qy   4165 ggcaaagtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgc 4224
Db   4059 gctaacgtacaaaccacggtaaatctcacggtgttgcaacaaccatttggacgttattgg 4118

Qy   4225 agtgcggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgcc 4284
Db   4119 caaaaagaagtgggattaaaggcagaaattttacatttccaaatttccgcttttatctca 4178

Qy   4285 gccgccaaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgc 4344
Db   4179 aaatctcaaggttcacaactcggcaaacagcaaaatgtgggcgtgaaattgggctatcgt 4238

Qy   4345 tggtaa 4350
Db   4239 tggtaa 4244
```

RESULT 12
AAZ12250
ID AAZ12250 standard; DNA; 2991 BP.
XX
AC AAZ12250;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria meningitidis partial ORF1 sequence.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT misc_feature 1316..1317
FT /*tag= a
FT /note= "826 unspecified nucleotides are present
FT between these bases"
FT misc_feature 3005..3006
FT /*tag= a
FT /note= "538 unspecified nucleotides are present
FT between these bases"
XX
PN WO9924578-A2.
XX
PD 20-MAY-1999.
XX
PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR P-PSDB; AAY38822.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 9; Page 358-359; 524pp; English.

XX
CC Nucleotide sequences AAZ11972-Z12358 represent open reading frames
CC (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode
CC antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their
CC fragments, their nucleic acids and antibodies are used for diagnosis,
CC prevention (as vaccines) or treatment of Neisseria infections,
CC such as meningitis, septicaemia and gonorrhea. Both organisms
CC are closely related. Fragments of the nucleic acids are useful
CC as hybridisation probes and antisense reagents.
XX
SQ Sequence 2991 BP; 840 A; 775 C; 694 G; 616 T; 66 other;

Query Match 24.1%; Score 1048.2; DB 20; Length 2991;
Best Local Similarity 73.4%; Pred. No. 2.7e-287;
Matches 1580; Conservative 7; Mismatches 82; Indels 485; Gaps 3;

```
Qy   2207 ccgacgataaagtgattgcttcattgactaagacngacntnagcggcantgtnagnctnn 2266
          | |  |||||||||| |||||||||||||||||| ||| | ||||||| |||    ||
Db   1313 ctgnngataaagtgactgcttcattgactaagaccgacatcagcggcaatgtcgatcttg 1372

Qy   2267 ccnatnacgntnnttnaaanctcncngggcntgcnncactnaanggcaatcttagtgcaa 2326
          || || ||| |  || ||| ||| | |||| |||  |||| || ||||||||||||||||
Db   1373 ccgatcacgctcatttaaatctcacagggcttgccacactcaacggcaatcttagtgcaa 1432

Qy   2327 atggcgatacacgttatacagtcagccacaacgccacccaaaacggcaaccttagcctcg 2386
          ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db   1433 atggcgatacacgttatacagtcagccacaacgccacccaaaacggcaaccktagcctcg 1492

Qy   2387 tgggcaatgcccaagcaacatttaatcaagccacattaaacggcaacncatcggnttcgg 2446
          ||| :|||||||||||||||||||||||||||||||||||||||||| |||||| |||||
Db   1493 tggnsaatgcccaagcaacatttaatcaagccacattaaacggcaacacatcggcttcgg 1552

Qy   2447 gcaatgcttcatttaatctaagcaacaacgccgcacaaaacggcagtctgacgctttccg 2506
          |||||||||||||||||||| || |||||| ||||||||||||||||||||||||||||
Db   1553 gcaatgcttcatttaatctaagcgaccacgccgtacaaaacggcagtctgacgctttccg 1612

Qy   2507 acaacgctaaggcaaacgtaagccattccgcactcaacggcaatgtctccctagccgata 2566
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1613 gcaacgctaaggcaaacgtaagccattccgcactcaacggtaatgtctccctagccgata 1672

Qy   2567 aggcagtattccattttgaaaacagccgctttaccggacaactcagcggcagcaaggana 2626
          |||||||||||||||||||| |||||||||||||||||||| |||||||| ||||||| |
Db   1673 aggcagtattccattttgaaagcagccgctttaccggacaaatcagcggcggcaaggata 1732

Qy   2627 cagcattacacttaaaagacagcgaatggacgctgccgtcaggcacggaattaggcaatt 2686
          | |||||||||||||||||||||||||||||||||||||||||| :|||||||||||||
Db   1733 cggcattacacttaaaagacagcgaatggacgctgccgtcaggarcggaattaggcaatt 1792

Qy   2687 taaaccttgacaacgccaccattacactcaattccgcctatcgccacgatgctgcaggcg 2746
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db   1793 taaaccttgacaacgccaccattacactcaattccgcctatcgccacgatgcggcagggg 1852

Qy   2747 cgcaaaccggcagngtgtcagacacgccgcgc---------cgccgttcgcgccgttccc 2797
          |||||||||||||| | | |||| |||||||         ||||||||||||||||||
Db   1853 cgcaaaccggcagtgcgacagatgcgccgcgccgccgttcgcgccgttcgcgccgttccc 1912

Qy   2798 tattatccgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggtaaacg 2857
          ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 tattatmcgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggtaaacg 1972

Qy   2858 gcaaattgaacngtcaaggaacattccgctttatgtcggaactcttcggctaccgaagcg 2917
          |||||||||| |||| ||||||||||||||||||||||||||||||||||||| ||||
Db   1973 gcaaattgaacggtcagggaacattccgctttatgtcggaactcttcggctaccgcagcg 2032

Qy   2918 acaaattgaagctggcggaaagttccgaaggnacttacaccttggcggtcaacaataccg 2977
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2033 acaaattgaagctggcggaaagttccgaaggcacttacaccttggcggtcaacaataccg 2092

Qy   2978 gcaacgaacccgtaagcctcgatcaattgacggtagtggaagggaaagacaacaaaccgc 3037
          |||||||||| | |||||||| ||||||||||||||||||| ||||||||||||||||
Db   2093 gcaacgaacctgcaagcctcgaacaattgacggtagtggaaggaaaagacaacaaaccgc 2152

Qy   3038 tgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgccggcgcgtggcgtt 3097
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||  ||
Db   2153 tgtccgaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcaggcgcgtggnntt 2212

Qy   3098 accaactcatccgcaaagacggcgagttccgcctgcataatccggtcaaagaacaagagc 3157
          |
Db   2213 a---------------------------------------------------------- 2213

Qy   3158 tttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaaa 3217

Db   2214 ---------------------------------------------------------- 2213

Qy   3218 gccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccg 3277

Db   2214 ---------------------------------------------------------- 2213

Qy   3278 aaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaa 3337

Db   2214 ---------------------------------------------------------- 2213

Qy   3338 aacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggc 3397

Db   2214 ---------------------------------------------------------- 2213

Qy   3398 cggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagc 3457

Db   2214 ---------------------------------------------------------- 2213

Qy   3458 cccaaccgcaacctcaaccccaaccgcagcgcgacctgatnagccgttatgccaatagcg 3517

Db   2214 ----------------------------------------------------------- 2213

Qy   3518 gtttgagtgaattttccgccacgctcaacagcgttttcgccgtacaggacgaattggacc 3577
                                                                  ||||
Db   2214 ----------------------------------------------------------gacc 2217

Qy   3578 gcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatccggnacaccaaac 3637
          |||| |||||||||||||||||||||| |||||||||||| |||||||| |||||||||
Db   2218 gcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatccgggacaccaaac 2277

Qy   3638 actaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggta 3697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 actaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggta 2337

Qy   3698 tgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaa 3757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2338 tgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaa 2397

Qy   3758 acancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgttttcgggc 3817
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2398 acaccttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgttttcgggc 2457

Qy   3818 aatacggcatcggcaggttcgacatcggcatcag-cacgggcgcgggttttagcagcggc 3876
          |||||||||||| ||||||| |||||||||||| | ||||||||||||||||||||||
Db   2458 aatacggcatcgacaggttctacatcggcatcagncgcgggcgcgggttttagcagcggc 2517

Qy   3877 antctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcattcag 3936
          |  || ||||||||||||||||:::|||:||||||||||||||||||||||||||||||
Db   2518 agcctttcagacggcatcggagsmaaawtccgccgccgcgtgctgcattacggcattcag 2577

Qy   3937 gcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaacgcgc 3996
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   2578 gcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatcggcgcaacgcgc 2637

Qy   3997 tatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccggtctt 4056
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   2638 tatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccggcctt 2697

Qy   4057 gcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaaccggcgcaacac 4116
          || |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Db   2698 gcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaaccggcgcaacac 2757

Qy   4117 atntccatcacnccttatttnagcctgtcctataccgatgccgcttcgggcaaagtccga 4176
```

```
Db    2758 atttccatcacgccttatttgagcctgtcctataccgatgccgcttcgggcaaagtccga 2817

Qy    4177 acacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgcagtgcggaatgg 4236
Db    2818 acacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgcagtgcggaatgg 2877

Qy    4237 ggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgccgccgccaaaggn 4296
Db    2878 ggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgccgccgccaaaggc 2937

Qy    4297 ccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggtaa 4350
Db    2938 ccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggtaa 2991
```

RESULT 13
AAA81401
ID AAA81401 standard; DNA; 891 BP.
XX
AC AAA81401;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis MenB internal polynucleotide sequence ORF number 76.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Disclosure; Page 219; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 891 BP; 255 A; 258 C; 202 G; 171 T; 5 other;

Query Match 18.2%; Score 790.8; DB 21; Length 891;
Best Local Similarity 92.4%; Pred. No. 2.4e-214;
Matches 823; Conservative 4; Mismatches 55; Indels 9; Gaps 1;

```
Qy    2212 gataaagtgattgcttcattgactaagacngacntnagcggcantgtnagnctnnccnat 2271
Db       1 gataaagtgactgcttcattgactaagaccgacatcagcggcaatgtcgatcttgccgat 60

Qy    2272 nacgntnnttnaaanctcncngggcntgcnncactnaanggcaatcttagtgcaaatggc 2331
Db      61 cacgctcatttaaatctcacagggcttgccacactcaacggcaatcttagtgcaaatggc 120

Qy    2332 gatacacgttatacagtcagccacaacgccacccaaaacggcaaccttagcctcgtgggc 2391
Db     121 gatacacgttatacagtcagccacaacgccacccaaaacggcaaccktagcctcgtggns 180

Qy    2392 aatgcccaagcaacatttaatcaagccacattaaacggcaacncatcggnttcgggcaat 2451
Db     181 aatgcccaagcaacatttaatcaagccacattaaacggcaacacatcggcttcgggcaat 240

Qy    2452 gcttcatttaatctaagcaacaacgccgcacaaaacggcagtctgacgctttccgacaac 2511
Db     241 gcttcatttaatctaagcgaccacgccgtacaaaacggcagtctgacgctttccggcaac 300

Qy    2512 gctaaggcaaacgtaagccattccgcactcaacggcaatgtctccctagccgataaggca 2571
Db     301 gctaaggcaaacgtaagccattccgcactcaacggtaatgtctccctagccgataaggca 360

Qy    2572 gtattccattttgaaaacagccgctttaccggacaactcagcggcagcaaggganacagca 2631
Db     361 gtattccattttgaaagcagccgctttaccggacaaatcagcggcggcaaggatacggca 420

Qy    2632 ttacacttaaaagacagcgaatggacgctgccgtcaggcacggaattaggcaatttaaac 2691
Db     421 ttacacttaaaagacagcgaatggacgctgccgtcaggarcggaattaggcaatttaaac 480

Qy    2692 cttgacaacgccaccattacactcaattccgcctatcgccacgatgctgcaggcgcgcaa 2751
Db     481 cttgacaacgccaccattacactcaattccgcctatcgccacgatgcggcaggggcgcaa 540

Qy    2752 accggcagngtgtcagacacgccgcgc---------cgccgttcgcgccgttccctatta 2802
Db     541 accggcagtgcgacagatgcgccgcgccgccgttcgcgccgttcgcgccgttccctatta 600

Qy    2803 tccgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggtaaacggcaaa 2862
Db     601 tmcgttacaccgccaacttcggtagaatcccgtttcaacacgctgacggtaaacggcaaa 660

Qy    2863 ttgaacngtcaaggaacattccgctttatgtcggaactcttcggctaccgaagcgacaaa 2922
Db     661 ttgaacggtcagggaacattccgctttatgtcggaactcttcggctaccgcagcgacaaa 720

Qy    2923 ttgaagctggcggaaagttccgaaggnacttacaccttggcggtcaacaataccggcaac 2982
Db     721 ttgaagctggcggaaagttccgaaggcacttacaccttggcggtcaacaataccggcaac 780

Qy    2983 gaacccgtaagcctcgatcaattgacggtagtggaagggaaagacaacaaaccgctgtcc 3042
Db     781 gaacctgcaagcctcgaacaattgacggtagtggaaggaaaagacaacaaaccgctgtcc 840

Qy    3043 gaaaaccttaatttcaccctgcaaaacgaacacgtcgatgccggcgcgtgg 3093
```

```
Db    841 gaaaaccttaatttcaccctgcaaaacgaacacgtcgatgcaggcgcgtgg 891


RESULT  14
AAA81402
ID   AAA81402 standard; DNA; 781 BP.
XX
AC   AAA81402;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis MenB 3' polynucleotide sequence ORF number 76.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Disclosure; Page 219; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 781 BP; 183 A; 256 C; 201 G; 137 T; 4 other;

  Query Match             17.0%;  Score 739.4;  DB 21;  Length 781;
  Best Local Similarity   96.5%;  Pred. No. 9.4e-200;

  Matches  754;  Conservative    3;  Mismatches   23;  Indels    1;  Gaps    1;

Qy   3571 ttggaccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatccggnac 3630
           || |||||||| |||||||||||||||||||||| |||||||||||| |||||||| ||
Db      1 ttagaccgcgtatttgccgaagaccgccgcaacgccgtttggacaagcggcatccgggac 60

Qy   3631 accaaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaa 3690
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 accaaaacactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaa 120

Qy   3691 atcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccgg 3750
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 atcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccgg 180

Qy   3751 accgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtt 3810
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 accgaaaacaccttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtt 240

Qy   3811 ttcgggcaatacggcatcggcaggttcgacatcggcatcag-cacgggcgcgggttttag 3869
           |||||||||||||||||| ||||||| |||||||||||| | ||||||||||||||||
Db    241 ttcgggcaatacggcatcgacaggttctacatcggcatcagncgcgggcgcgggttttag 300

Qy   3870 cagcggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacgg 3929
           |||||||| || ||||||||||||||::|||:|||||||||||||||||||||||||||
Db    301 cagcggcagcctttcagacggcatcggagsmaaawtccgccgccgcgtgctgcattacgg 360

Qy   3930 cattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgc 3989
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 cattcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgcacatcggcgc 420

Qy   3990 aacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccc 4049
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 aacgcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccc 480

Qy   4050 cggtcttgcgttcaaccgntaccgngcgggcattaaggcagattattcattcaaaccggc 4109
           ||| ||||| |||||||| ||||| ||||||||||||||||||||||||||||||||||
Db    481 cggccttgcattcaaccgctaccgcgcgggcattaaggcagattattcattcaaaccggc 540

Qy   4110 gcaacacatntccatcacncctttatttnagcctgtcctataccgatgccgcttcgggcaa 4169
           ||||||||| |||||||| |||||||| ||||||||||||||||||||||||||||||||
Db    541 gcaacacatttccatcacgccttatttgagcctgtcctataccgatgccgcttcgggcaa 600

Qy   4170 agtccgaacacgcgtcaataccgcngtattggctcaggatttcggcaaaacccgcagtgc 4229
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    601 agtccgaacacgcgtcaataccgccgtattggctcaggatttcggcaaaacccgcagtgc 660

Qy   4230 ggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccntccacgctgccgccgc 4289
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    661 ggaatggggcgtaaacgccgaaatcaaaggtttcacgctgtccctccacgctgccgccgc 720

Qy   4290 caaaggnccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggta 4349
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 caaaggcccgcaactggaagcgcaacacagcgcgggcatcaaattaggctaccgctggta 780

Qy   4350 a 4350
           |
Db    781 a 781


RESULT  15
AAA81400
ID   AAA81400 standard; DNA; 1312 BP.
XX
AC   AAA81400;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis MenB 5' polynucleotide sequence ORF number 76.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
```

OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Disclosure; Page 218-219; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 1312 BP; 402 A; 261 C; 291 G; 308 T; 50 other;

Query Match 15.9%; Score 690.8; DB 21; Length 1312;
Best Local Similarity 72.2%; Pred. No. 8.6e-186;
Matches 1000; Conservative 4; Mismatches 297; Indels 84; Gaps 8;

```
Qy      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggccgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      1 atgaaaacaaccgacaaacggacaaccgaaacacaccgcaaagccccgaaaaccggtcgc 60

Qy     61 atccgcttctcgcctgcttacttagccatatgcctgtcgttcggcattcttccccaagct 120
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 atccgcttctcg-ctgcttacttagccatatgcctgtcgttcggcattcttccccaagcc 119

Qy    121 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 tgggcgggacacacttatttcggcatcaactaccaatactatcgcgactttgccgaaaat 179

Qy    181 aaaggcaagtttgcagtcggggcgaaagatattgaggtntacaacaaaaaaggggagttg 240
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    180 aaaggcaagtttgcagtcggggcgaaagatattgaggtttacaacaaaaaaggggagttg 239

Qy    241 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 gtcggcaaatcaatgacaaaagccccgatgattgatttttctgtggtgtcgcgtaacggc 299

Qy    301 gtggcggcattggtgggcg-atcaatatattgtgagcgtggcacataacggcggctataa 359
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    300 gtggcggcattggtgggcgtatcaatatattgtgagcgtggcacataacggcggctataa 359

Qy    360 caacgttgattttggtgcggaaggaagna-atcccgatcagcaccgtttttcttaccaaa 418
          |||||||||||||||||||||||||||: | |||||||||| || ||::|| |||| |||
Db    360 caacgttgattttggtgcggaaggaakaatatcccgatcaacawcgwwttacttataaaa 419

Qy    419 ttgtgaaaagaaataattataagcctgacaattca---caccctacaacggcgattanc 475
          |||||||| | |||||||||||  | |  | |  |   || |||||   |||||||| |
Db    420 ttgtgaaacggaataattataaagcagggactaaaggccatccttatggcggcgattatc 479

Qy    476 atatgccgcgtttgcataaatttgtcacagatgcagaacctgtcgaaatgacgagtgaca 535
          ||||||||||||||||||||:|||||||||||||||||||| |||||||| ||| | |
Db    480 atatgccgcgtttgcataaatwtgtcacagatgcagaacctgttgaaatgaccagttata 539

Qy    536 tgaggggaatacctattccgataaagaaaaatatcccgagcgtgtccgcatcggctcag 595
          ||   |||   |  |||  |||| || | || || || || ||||| || || ||  |||
Db    540 tggatgggcggaaatatatcgatcaaaataattaccctgaccgtgttcgtattggggcag 599

Qy    596 gacaccactattggcgttatgatgatgacaaacacggcgatttatcctactccggcgcat 655
          |    || ||||||||  | ||||||  ||  | |
Db    600 gcaggcaatattggcgatctgatgaagatgagc---------------------------- 632

Qy    656 ggttaattggcggcaatacacatatgcagggttggggaaataatggcgtanttagtttga 715
                          |  |||  |  |   |   | || ||   | |  |
Db    633 ------------------ccaataaccgcgaaagttcatatcatattgcaagt------- 667

Qy    716 gcggcgatgtgcgccatgccaacgactatggccctatgccgattgcaggtgcggcaggcg 775

Db    668 ----------nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 707

Qy    776 acagcggttcgccaatgtttatttatgacaaaacaaacaataaatggctgctcaacggag 835
              || || |||||||||||| |||||     |||  | || ||| |  |  || || |
Db    708 nnnnnggctcaccaatgtttatctatgatgcccaaaagcaaaagtggttaattaatgggg 767

Qy    836 ttttacaaaccggctacccttattccggcagggaaaacggtttccagctgatacgcaaag 895
          |  || ||||| ||| |||| |||   ||  |     ||  || |||||||| | || ||||
Db    768 tattgcaaacgggcaacccctatataggaaaaagcaatggcttccagctggttcgtaaag 827

Qy    896 attggttctacgatgacatttacagaggcgatacacataccgtctntttttgaaccgcgca 955
          |||||||||| |||||  || |     ||  |||||  ||| | || |  |  |||||  ||
Db    828 attggttctatgatgaaatctttgctggagatacccattcagtattctacgaaccacgtc 887

Qy    956 gtaacggacatttttccttacatccaacaacaacggtacgggtacggtaacagaaacca 1015
            ||  ||  | |  ||  ||||    | | || || || || || ||  |  | |  |
Db    888 aaaatgggaaatactcttttaacgacgataataatggcacaggaaaaatcaatgccaaac 947

Qy   1016 acgaaaaggtntccaatccaaa---gcttaaagtacagacagtccgactgtttgacgaat 1072
          | ||| |    ||    || ||     | |||  ||  || || | | |||||  | |  |
Db    948 atgaacacaattctctgcctaatagattaaaaacacgaaccgttcaattgtttaatgttt 1007

Qy   1073 ctttgaatgaaactgataaagaaccagttt---acgcggcaggggtgttaatcagtacc 1129
          ||||    || || |  | ||||||  ||||    | || |||||  |||||  ||    || |
Db   1008 ctttatccgagacagcaagagaacctgtttatcatgctgcaggtggtgtcaacagttatc 1067

Qy   1130 gtccaaggttaaacaacggtgaaaaccttttcttttatcgattacggcaacggcaaactca 1189
          | || ||  | || || || |||||  |||| |||||  ||  | || || ||| || | |
Db   1068 gacccagactgaataatggagaaaatatttcctttattgacgaaggaaaaggcgaattga 1127

Qy   1190 tcttatcaaacaacatcaaccaaggcgcgggcggtttgtattttgaaggtgattttacgg 1249
          |  |  | | |||||||||  |||||  || || || ||  |||||  ||||  ||||||||||
Db   1128 tacttaccagcaacatcaatcaaggtgctggaggattatatttccaaggagattttacgg 1187

Qy   1250 tctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatcagtgaagacagta 1309
          |||||||||||| ||||||||  ||||||||||||||||||||||||||||||||||||||
Db   1188 tctcgcctgaaaataacgaaacttggcaaggcgcgggcgttcatatcagtgaagacagta 1247
```

```
Qy   1310 ccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaaatcggcaaaggca 1369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1248 ccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaaatcggcaaaggca 1307

Qy   1370 cgctg 1374
          |||||
Db   1308 cgctg 1312


RESULT 16
AAT88142
ID   AAT88142 standard; DNA; 969 BP.
XX
AC   AAT88142;
XX
DT   11-MAY-1998 (first entry)
XX
DE   H. influenzae Hap autotransporter membrane integration region DNA.
XX
KW   Hap gene; autotransporter; Gram-negative bacteria; diagnostic; therapy;
KW   surface presented polypeptide; ss.
XX
OS   Haemophilus influenzae.
XX
PN   WO9735022-A1.
XX
PD   25-SEP-1997.
XX
PF   15-MAR-1996;   96WO-EP01130.
XX
PR   15-MAR-1996;   96WO-EP01130.
XX
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI   Jose J, Maurer J, Meyer TF;
XX
DR   WPI; 1997-480227/44.
DR   P-PSDB; AAW27705.
XX
PT   Presentation of peptide(s) on surface of Gram-negative bacteria -
PT   via transformation with vector encoding signal peptide, presented
PT   peptide and transporter domain of auto-transporter, producing
PT   peptide libraries for epitope mapping
XX
PS   Claim 8; Fig 9; 84pp; German.
XX
CC   This DNA sequence encodes the H. influenzae Hap autotransporter membrane
CC   integration region. This region is involved in a novel method which
CC   allows the presentation of stable fusion polypeptides on the surface of
CC   Gram-negative bacteria which can be released into the surrounding media.
CC   The method can be used to produce a variegated population of
CC   surface-presented polypeptides, so that bacteria expressing polypeptides
CC   with particular properties can be identified and simultaneously selected,
CC   e.g. for epitope mapping or selection of ligands with the highest
CC   affinity for antibodies, major histocompatibility complex (MHC) molecules
CC   or other components of the immune system. Selected polypeptides can be
CC   used diagnostically, e.g. to screen sera or antibody banks, and
CC   polypeptide expressing cells may be used as live vaccines. They may also
CC   be used therapeutically, e.g. when the polypeptide is an antibody, to
CC   remove or concentrate pollutants, inactivate toxins, prepare and process
CC   food, prepare washing compositions and label cells. Selected bacteria can
CC   be stored, reproduced and replicated on a large scale as individual
CC   clones.
XX
SQ   Sequence 969 BP; 313 A; 178 C; 212 G; 266 T; 0 other;


  Query Match             4.7%;  Score 206;  DB 18;  Length 969;
  Best Local Similarity   52.6%;  Pred. No. 6.8e-48;
  Matches  453;  Conservative    0;  Mismatches  405;  Indels    3;  Gaps    1;

Qy   3490 gacctgatnagccgttatgccaatagcggtttgagtgaattttccgccacgctcaacagc 3549
          ||| |||| |||||||||  | |||||  |  ||    ||||| || || ||  | || ||

Db    109 gacttgatcagccgttattcaaatagtgcgttatcagaattatctgcaacagtaaatagt 168

Qy   3550 gttttcgccgtacaggacgaattggaccgcgtgtttgccgaagaccgccgcaacgcngtt 3609
          |  |  | || || || ||||| || ||  | ||||  ||  |    |     || ||
Db    169 atgctttctgttcaagatgaattagatcgtctttttgtagatcaagcacaatctgccgtg 228

Qy   3610 tggacaagcngcatccggnacaccaaacactaccgttcgcaagatttccgcgcctaccgc 3669
          |||||||    |   | | | |  | || |||   |||  | |  ||||| || || |
Db    229 tggacaaatatcgcacaggataaaagacgctatgattctgatgcgttccgtgcttatcag 288

Qy   3670 caac---aaaccgacctgcgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtc 3726
          || |   ||||  || | || |||||  ||  |||| |||  | | |  |  || ||  |
Db    289 cagcagaaaacgaacttacgtcaaattggggtgcaaaaagccttagctaatggacgaatt 348

Qy   3727 ggcatcctgttttcgcacaaccggaccgaaaacancttcgacgacggcatcggcaactcg 3786
          ||     | || ||||| | |||  | || || | ||| || ||     |    ||
Db    349 ggggcagttttctcgcatagccgttcagataataccttgatgaacaggttaaaaatcac 408

Qy   3787 gcacggcttgcccacggcgccgttttcgggcaatacggcatcggcaggttcgacatcggc 3846
          ||     |  |        | | ||| |  |||||       |||  ||  |  | ||
Db    409 gcgacattaacgatgatgtcgggttttgcccaatatcaatggggcgatttacaatttggt 468

Qy   3847 atcagcacgggcgcgggttttagcagcggcantctntcagacggcatcggaggcaaaatc 3906
          | | |  |||  ||||  | ||     | |   |  | || |       |  | |||||
Db    469 gtaaacgtgggaacgggaatcagtgcgagtaaaatggctgaagaacaaagccgaaaaatt 528

Qy   3907 cgccgccgcgtgctgcattacggcattcaggcacgataccgcgccggtttcggcggattc 3966
          |  ||    | | |  |||| ||| |  | ||| | || |    | |||| ||  |||
Db    529 catcgaaaagcgataaattatggcgtgaatgcaagttatcagttccgtttagggcaattg 588

Qy   3967 ggcatcgaaccgtacatcggcgcaacgcgctatttcgtccaaaaagcggattaccgctac 4026
          |||||  | || ||  | || |  |  ||||||||  |  ||   |  |||| |  |
Db    589 ggcattcagccttattttggagttaatcgctattttattgaacgtgaaaattatcaatct 648

Qy   4027 gaaaacgtcaatatcgccacccccggtcttgcgttcaaccgntaccgngcgggcattaag 4086
          ||  | || |   |    || ||  | ||||| || || || ||    || ||||||
Db    649 gaggaagtgagagtgaaaacgcctagccttgcatttaatcgctataatgctggcattcga 708

Qy   4087 gcagattattcattcaaaccggcgcaacacatntccatcacnccttatttnagcctgtcc 4146
          |  |||||| |||| |  ||| |  |  | ||   | | |  ||||||||   | |
Db    709 gttgattatacatttactccgacagataatatcagcgttaagccttatttcttcgtcaat 768

Qy   4147 tataccgatgccgcttcgggcaaagtccgaacacgcgtcaataccgcngtattggctcag 4206
          |||   ||||  |   |  || || | |||   || |||  | | || |||   ||
Db    769 tatgttgatgtttcaaacgctaacgtacaaaccacggtaaatctcacggtgttgcaacaa 828

Qy   4207 gatttcggcaaaacccgcagtgcggaatggggcgtaaacgccgaaatcaaaggtttcacg 4266
             || ||        |       |||  |||  |||| || |||||   |  ||||
Db    829 ccatttggacgttattggcaaaaagaagtgggattaaaggcagaaattttacatttccaa 888

Qy   4267 ctgtccntccacgctgccgccgccaaaggnccgcaactggaagcgcaacacagcgcgggc 4326
          | |||          |     |  ||||  | ||||| |     || || |  | ||||
Db    889 atttccgcttttatctcaaaatctcaaggttcacaactcggcaaacagcaaaatgtgggc 948

Qy   4327 atcaaattaggctaccgctgg 4347
          | ||||| ||||| || |||
Db    949 gtgaaattgggctatcgttgg 969


RESULT 17
AAA81520/c
ID   AAA81520 standard; DNA; 30078 BP.
XX
AC   AAA81520;
XX
DT   04-DEC-2000 (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_67 SEQ ID NO:67.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
```

```
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1411-1420; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 30078 BP; 6365 A; 7264 C; 8488 G; 7960 T; 1 other;

  Query Match             4.4%;  Score 191.4;  DB 21;  Length 30078;
  Best Local Similarity   58.8%;  Pred. No. 7.3e-43;
  Matches  346;  Conservative    0;  Mismatches  239;  Indels    3;  Gaps    1;

Qy   1161 ttttatcgattacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggg 1220
           | ||| |  | |   ||||||||   |  | ||     | ||||||||||||||||||||
Db  22246 TGTTACCTTTGAAAACAACGGCACTTTGGTATTGGATCAAAACATCAACCAAGGCGCGGG 22187

Qy   1221 cggtttgtattttgaaggtgattttacggtctcgcctgaaaacaacga---aacgtggca 1277
           |||| ||| |||  |||| ||||  || |||     |  ||| || ||    || ||||
Db  22186 CGGTCTGTTTTTCAAAGGCGATTACACAGTCAAAGGTATAAATAATGACATCACTTGGCT 22127

Qy   1278 aggcgcgggcgttcatatcagtgaagacagtaccgttacttggaaagtaaacggcgtggc 1337
           ||| |||||  || || |    || | ||  |  ||   |||| |||| ||      |
Db  22126 AGGTGCGGGGATTGATGTTGCCGACGGCAAAAAAGTCGTTTGGCAAGTCAAAAATCCGAA 22067

Qy   1338 aaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaa 1397
             |||| |  || | ||||||||||||||||||  |  |  |  | | || ||| |  ||
Db  22066 TGGCGACAGATTGGCAAAAATCGGCAAAGGCACATTAGAAATAAACGGCACAGGCGTTAA 22007

Qy   1398 ccaaggctcgatcagcgtgggcgacggtacagtcattttggatcagcaggcagacgataa 1457
           ||||||     | |  || ||||||||||| || ||| || |||| || || || |   |
Db  22006 CCAAGGGCAATTAAAAGTCGGCGACGGTACGGTTATTCTGAATCAACAAGCCGATGCCGA 21947

Qy   1458 aggcaaaaaacaagcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaact 1517
             |||   || || ||     || ||||| |  |||||||  | |||||  ||  | |
Db  21946 CAAAAAAGTCCAGGCTTTCTCCCAAGTCGGCATTGTCAGCGGACGCGGTACATTGGTATT 21887

Qy   1518 gaatgccgataatcagttcaaccccgacaaactctatttcggctttcgcggcggacgttt 1577
           |||      |||||| | || ||||| || || |||||||| || || ||||| |||||
Db  21886 AAATAGTTCAAATCAGATTAATCCCGATAACCTATATTTCGGTTTCCGTGGCGGTCGTTT 21827

Qy   1578 ggatttaaacgggcattcgctttcgttccaccgtattcaaaataccgatgaaggggcgat 1637
           ||||   || ||  ||     |  | ||  | |  || |  ||    |||||||| |||
Db  21826 GGATGCCAATGGCAATGACTTGACTTTTGAACACATCCGCAACGTGGATGAAGGCGCGCG 21767

Qy   1638 gattgncnatcataatgccacaacaacatccaccgttaccattacagggaatgaaagtat 1697
           |||| | | || ||  |        | |||||  | ||  | || || ||       ||
Db  21766 CATTGTCAACCACAACACAGGCCACGCCTCCACAATCACGCTAACGGGTAAATCTTTGAT 21707

Qy   1698 tacacaaccgagtggtaagaatatcaatagacttaattacagcaaaga 1745
           |||  | || |    |    |||  ||    || |  | |   | ||
Db  21706 TACCGATCCCAAAACCATCTCTATTCATTATATTCAAAATAATGATGA 21659


RESULT  18
AAF21608/c
ID   AAF21608 standard; DNA; 349980 BP.
XX
AC   AAF21608;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:109.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning  of AAF21607, the last 49980 bp of AAF21607 are repeated at
```

```
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 82523 A; 82940 C; 96712 G; 87805 T; 0 other;


  Query Match             4.4%;  Score 191.4;  DB 21;  Length 349980;
  Best Local Similarity   58.8%;  Pred. No. 3.1e-42;
  Matches  346;  Conservative    0;  Mismatches  239;  Indels    3;  Gaps    1;

Qy    1161 ttttatcgattacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcggg 1220
           | ||| |  | |   |||||||||   |  | ||     | ||||||||||||||||||||
Db  130048 TGTTACCTTTGAAAACAACGGCACTTTGGTATTGGATCAAAACATCAACCAAGGCGCGGG 129989

Qy    1221 cggtttgtattttgaaggtgattttacggtctcgcctgaaaacaacga---aacgtggca 1277
           |||| ||| |||  |||| ||||  || |||     |  ||| || ||    || ||||
Db  129988 CGGTCTGTTTTTCAAAGGCGATTACACAGTCAAAGGTATAAATAATGACATCACTTGGCT 129929

Qy    1278 aggcgcgggcgttcatatcagtgaagacagtaccgttacttggaaagtaaacggcgtggc 1337
           ||| |||||  || || |    || | ||  |  ||   |||| |||| ||      |
Db  129928 AGGTGCGGGGATTGATGTTGCCGACGGCAAAAAAGTCGTTTGGCAAGTCAAAAATCCGAA 129869

Qy    1338 aaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaa 1397
            ||||  |  || | ||||||||||||||||  |  |  |  | | || ||| |  ||
Db  129868 TGGCGACAGATTGGCAAAAATCGGCAAAGGCACATTAGAAATAAACGGCACAGGCGTTAA 129809

Qy    1398 ccaaggctcgatcagcgtgggcgacggtacagtcatttggatcagcaggcagacgataa 1457
           ||||||     | |  || |||||||||||| || ||| || |||| || || || |  |
Db  129808 CCAAGGGCAATTAAAAGTCGGCGACGGTACGGTTATTCTGAATCAACAAGCCGATGCCGA 129749

Qy    1458 aggcaaaaaacaagcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaact 1517
              |||   || || ||     || ||||| |  |||||||  | |||||  ||  | |
Db  129748 CAAAAAAGTCCAGGCTTTCTCCCAAGTCGGCATTGTCAGCGGACGCGGTACATTGGTATT 129689

Qy    1518 gaatgccgataatcagttcaaccccgacaaactctatttcggctttcgcggcggacgttt 1577
            |||      |||||| | || ||||| || || |||||||| || || ||||| |||||
Db  129688 AAATAGTTCAAATCAGATTAATCCCGATAACCTATATTTCGGTTTCCGTGGCGGTCGTTT 129629

Qy    1578 ggatttaaacgggcattcgctttcgttccaccgtattcaaaataccgatgaaggggcgat 1637
           ||||   || ||  ||    |  | ||  | |  || |  ||    |||||||| |||
Db  129628 GGATGCCAATGGCAATGACTTGACTTTTGAACACATCCGCAACGTGGATGAAGGCGCGCG 129569

Qy    1638 gattgncnatcataatgccacaacaacatccaccgttaccattacagggaatgaaagtat 1697
            |||| | | || ||  |        | |||||  | ||  | || || ||       ||
Db  129568 CATTGTCAACCACAACACAGGCCACGCCTCCACAATCACGCTAACGGGTAAATCTTTGAT 129509

Qy    1698 tacacaaccgagtggtaagaatatcaatagacttaattacagcaaaga 1745
           |||  | || |     |    |||  ||    || |  | |   | ||
Db  129508 TACCGATCCCAAAACCATCTCTATTCATTATATTCAAAATAATGATGA 129461


RESULT  19
AAZ54326
ID   AAZ54326 standard; DNA; 4296 BP.
XX
AC   AAZ54326;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 759 partial DNA sequence SEQ ID NO:2601.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; AAY75564.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1232-1234; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 4296 BP; 1192 A; 1470 C; 959 G; 675 T; 0 other;


  Query Match             4.1%;  Score 177.4;  DB 21;  Length 4296;
  Best Local Similarity   60.0%;  Pred. No. 2.3e-39;
  Matches  311;  Conservative    0;  Mismatches  204;  Indels    3;  Gaps    1;

Qy    1171 tacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtat 1230
           | || ||||   | ||| ||  |  || |||| ||||||||||||||| ||||  ||| |
Db    1060 ttcgacaacaaaacactgatgctggcagacaatatcaaccaaggcgcaggcgcattgcag 1119

Qy    1231 tttgaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgtt 1290
           || ||  |  | || || |||     | ||||| ||   || |||||||| || ||||||
Db    1120 ttcgacagcaacttcaccgtcgtcggtaaaaaccac---acatggcaaggtgcaggcgtt 1176

Qy    1291 catatcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctg 1350
              |   || | ||    |||    ||| |||| | |  |     | |||||| ||
Db    1177 atcgtagccgacggcaaacgcgtcttctggcaagtcagcaaccccaaaggcgaccggctc 1236

Qy    1351 tccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatc 1410
           ||||||  | |||  |||||||||   ||   | |   ||||    ||||| |||   |||
Db    1237 tccaaactgggcgcaggcacgcttatcgccaacggacaaggcatcaaccagggcgacatc 1296
```

```
Qy   1411 agcgtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaa 1470
          ||| | || || || || ||| |  | |  ||  | || |      | ||||  ||||||
Db   1297 agcatcggggaaggcactgtcgtactcgcccaaaaagctgcttcagacggcagcaaacaa 1356

Qy   1471 gcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataat 1530
          || || |   || ||||| |   ||||||||||||| ||||     ||    | |    |
Db   1357 gcattcaaccaagtcggcatcaccagcggcaggggcacggccgtcctcgccgacagccag 1416

Qy   1531 cagttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacggg 1590
          ||  |||| ||||| || ||||||||||||||  | ||||||||  | ||  | |||||
Db   1417 caaatcaaacccgaaaacctctatttcggcttcaggggcggacggctcgacctcaacggc 1476

Qy   1591 cattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
            |    ||| | ||   || ||| |   || | || |  || |||   || | | ||||
Db   1477 aacaaccttgcctttacccatatccgccatgcggacggcggcgcgcaaatcgtcaatcac 1536

Qy   1651 aatgccacaacaacatccaccgttaccattacagggaa 1688
          ||  |       | |  | ||  | ||  | || || ||
Db   1537 aaccctgaccaagccgcgacactgacgctgaccggcaa 1574
```

RESULT 20
AAA81472/c
ID AAA81472 standard; DNA; 16526 BP.
XX
AC AAA81472;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 466-471; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 2 other;

Query Match 4.1%; Score 177.4; DB 21; Length 16526;
Best Local Similarity 60.0%; Pred. No. 5e-39;
Matches 311; Conservative 0; Mismatches 204; Indels 3; Gaps 1;

```
Qy   1171 tacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtat 1230
          | || ||||  | ||| ||  |  || |||| |||||||||||||| ||||  ||| |
Db   7260 TTCGACAACAAAACACTGATGCTGGCAGACAATATCAACCAAGGCGCAGGCGCATTGCAG 7201

Qy   1231 tttgaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgtt 1290
          || ||  |  | || || |||     | ||||| ||   || |||||||| || ||||||
Db   7200 TTCGACAGCAACTTCACCGTCGTCGGTAAAAACCAC---ACATGGCAAGGTGCAGGCGTT 7144

Qy   1291 catatcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctg 1350
            |   || | ||   |||   ||| |||| | |  |     | |||||| ||
Db   7143 ATCGTAGCCGACGGCAAACGCGTCTTCTGGCAAGTCAGCAACCCCAAAGGCGACCGGCTC 7084

Qy   1351 tccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatc 1410
          |||||| | |||  |||||||||   ||   | |  ||||   |||||  |||  |||
Db   7083 TCCAAACTGGGCGCAGGCACGCTTATCGCCAACGGACAAGGCATCAACCAGGGCGACATC 7024

Qy   1411 agcgtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaa 1470
          ||| | || || || || ||| |  | |  ||  | || |      | ||||  ||||||
Db   7023 AGCATCGGGGAAGGCACTGTCGTACTCGCCCAAAAAGCTGCTTCAGACGGCAGCAAACAA 6964

Qy   1471 gcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataat 1530
          || || |   || ||||| |   ||||||||||||| ||||     ||    | |    |
Db   6963 GCATTCAACCAAGTCGGCATCACCAGCGGCAGGGGCACGGCCGTCCTCGCCGACAGCCAG 6904

Qy   1531 cagttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacggg 1590
          ||  |||| ||||| || ||||||||||||||  | ||||||||  | ||  | |||||
Db   6903 CAAATCAAACCCGAAAACCTCTATTTCGGCTTCAGGGGCGGACGGCTCGACCTCAACGGC 6844

Qy   1591 cattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
            |   ||| | ||   || ||| |   || | || |  || |||   || | | ||||
Db   6843 AACAACCTTGCCTTTACCCATATCCGCCATGCGGACGGCGGCGCGCAAATCGTCAATCAC 6784

Qy   1651 aatgccacaacaacatccaccgttaccattacagggaa 1688
          ||  |       | |  | ||  | ||  | || || ||
Db   6783 AACCCTGACCAAGCCGCGACACTGACGCTGACCGGCAA 6746
```

RESULT 21
AAF21613
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.

XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 other;

Query Match 4.1%; Score 177.4; DB 21; Length 172325;
Best Local Similarity 60.0%; Pred. No. 2e-38;
Matches 311; Conservative 0; Mismatches 204; Indels 3; Gaps 1;

```
Qy   1171 tacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtat 1230
          | || ||||   | ||| ||  |  || |||| ||||||||||||||| ||||  |.|| |
Db   8178 ttcgacaacaaaacactgatgctggcagacaatatcaaccaaggcgcaggcgcattgcag 8237

Qy   1231 tttgaaggtgatttttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgtt 1290
          || ||  |  | || || |||     | ||||| ||   || |||||||| || ||||||
Db   8238 ttcgacagcaacttcaccgtcgtcggtaaaaaccac---acatggcaaggtgcaggcgtt 8294

Qy   1291 catatcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctg 1350
               |    || | ||    |||    ||| |||| | |  |     |  |||||| ||
Db   8295 atcgtagccgacggcaaacgcgtcttctggcaagtcagcaaccccaaaggcgaccggctc 8354

Qy   1351 tccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatc 1410
          |||||| | |||  |||||||||   ||   | |   ||||    ||||| |||   |||
Db   8355 tccaaactgggcgcaggcacgcttatcgccaacggacaaggcatcaaccagggcgacatc 8414

Qy   1411 agcgtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaa 1470
          ||| | || || || || ||| |  | |  ||  | || |      | ||||  ||||||
Db   8415 agcatcggggaaggcactgtcgtactcgcccaaaaagctgcttcagacggcagcaaacaa 8474

Qy   1471 gcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataat 1530
          || || |   || ||||| |   |||||||||||| ||||     ||    | |    |
Db   8475 gcattcaaccaagtcggcatcaccagcggcaggggcacggccgtcctcgccgacagccag 8534

Qy   1531 cagttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacggg 1590
          ||  |||| ||||| || ||||||||||||||  | ||||||||  | ||  | |||||
Db   8535 caaatcaaacccgaaaacctctatttcggcttcaggggcggacggctcgacctcaacggc 8594

Qy   1591 cattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
            |    ||| | ||   || ||| |   || | || |  || |||   || | | ||||
Db   8595 aacaaccttgcctttacccatatccgccatgcggacggcggcgcgcaaatcgtcaatcac 8654

Qy   1651 aatgccacaacaacatccaccgttaccattacagggaa 1688
          || |      | |  | ||  | ||  | || || ||
Db   8655 aaccctgaccaagccgcgacactgacgctgaccggcaa 8692
```

RESULT 22
AAF21612
ID AAF21612 standard; DNA; 349980 BP.
XX
AC AAF21612;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection

```
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 other;


  Query Match             4.1%;  Score 177.4;  DB 21;  Length 349980;
  Best Local Similarity   60.0%;  Pred. No. 3e-38;
  Matches  311;  Conservative    0;  Mismatches  204;  Indels    3;  Gaps    1;

Qy   1171 tacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtat 1230
Db 308178 ttcgacaacaaaacactgatgctggcagacaatatcaaccaaggcgcaggcgcattgcag 308237

Qy   1231 tttgaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgtt 1290
Db 308238 ttcgacagcaacttcaccgtcgtcggtaaaaaccac---acatggcaaggtgcaggcgtt 308294

Qy   1291 catatcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctg 1350
Db 308295 atcgtagccgacggcaaacgcgtcttctggcaagtcagcaaccccaaaggcgaccggctc 308354

Qy   1351 tccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatc 1410
Db 308355 tccaaactgggcgcaggcacgcttatcgccaacggacaaggcatcaaccagggcgacatc 308414

Qy   1411 agcgtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaa 1470
Db 308415 agcatcggggaaggcactgtcgtactcgcccaaaaagctgcttcagacggcagcaaacaa 308474

Qy   1471 gcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataat 1530
Db 308475 gcattcaaccaagtcggcatcaccagcggcaggggcacggccgtcctcgccgacagccag 308534

Qy   1531 cagttcaacccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacggg 1590
Db 308535 caaatcaaacccgaaaacctctatttcggcttcaggggcggacggctcgacctcaacggc 308594

Qy   1591 cattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
Db 308595 aacaaccttgcctttacccatatccgccatgcggacggcggcgcgcaaatcgtcaatcac 308654

Qy   1651 aatgccacaacaacatccaccgttaccattacagggaa 1688
Db 308655 aaccctgaccaagccgcgacactgacgctgaccggcaa 308692


RESULT  23
AAQ06164
ID   AAQ06164 standard; DNA; 5091 BP.
XX
AC   AAQ06164;
XX
DT   31-JAN-1991  (first entry)
XX
DE   The iga gene, encoding IgA1 protease.
XX
KW   IgA1; vaccine; meningitis; gonnorhoea; allergies; ss.
XX
OS   Haemophilus influenzae.
XX
FH   Key             Location/Qualifiers
FT   CDS             262..4887
FT                   /*tag=  a
FT                   /product=IgA1 protease
FT   -10_signal      235..240
FT                   /*tag=  b
FT   -35_signal      211..215
FT                   /*tag=  c
FT   RBS             248..253
FT                   /*tag=  d
XX
PN   WO9011367-A.
XX
PD   04-OCT-1990.
XX
PF   16-MAR-1990;   90WO-DK00073.
XX
PR   17-MAR-1989;   89DK-0001308.
XX
PA   (KILI/) KILIAN M.
XX
PI   Kilian M,  Poulsen K;
XX
DR   WPI; 1990-320267/42.
DR   P-PSDB; AAR07304.
XX
PT   Immunoglobulin A1 protease prodn. - by cloning from
PT   microorganisms for immunisation against immunoglobulin A1
PT   protease producing bacteria
XX
PS   Disclosure; fig 3; 44pp; English.
XX
CC   This iga gene is from H.influenzae serotype b strain HK368.  On
CC   transformation of E.coli cells with a vector contg. this gene,
CC   immunoglobulin (Ig)A1 protease is produced which is useful in a
CC   vaccine for e.g. meningococcal meningitis, gonorrhoea or allergic
CC   diseases.
XX
SQ   Sequence 5091 BP; 1862 A; 873 C; 979 G; 1377 T; 0 other;


  Query Match             3.6%;  Score 158;  DB 11;  Length 5091;
  Best Local Similarity   57.7%;  Pred. No. 8.3e-34;
  Matches  319;  Conservative    0;  Mismatches  228;  Indels    6;  Gaps    2;

Qy   1161 ttttatcgattacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcggg 1220
Db   1410 tgttacatttgaagggagtggaacgcttaccttaaataataatatcgatcaaggtgcagg 1469

Qy   1221 cggtttgtattttgaaggtgattttacggtctcgc---ctgaaaacaacgaaacgtggca 1277
Db   1470 cggattattctttgaaggcgattatgaagttaaaggtacttctgataatactacttggaa 1529

Qy   1278 aggcgcgggcgttcatatcagtgaagacagtaccgttacttggaaagtaaacggcgtggc 1337
Db   1530 aggagcaggtgtctctgttgccgaaggaaaaactgtaacgtggaaagtgcataatcctca 1589

Qy   1338 aaacgaccgcctgtccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaa 1397
Db   1590 atatgatcgtttagcaaaaattggcaaagggacattaattgttgaaggaacaggagataa 1649

Qy   1398 ccaaggctcgatcagcgtgggcgacggtacagtcattttggatcagcaggcagacgataa 1457
Db   1650 taaaggttcgctaaaagtgggcgatggcaccgttatttttaaaacaacaaacaaatggttc 1709

Qy   1458 aggcaaaaaacaagcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaact 1517
Db   1710 ggg---acaacacgcttttgcttctgtagggattgtaagtggtcgctcaactcttgtgct 1766

Qy   1518 gaatgccgataatcagttcaaccccgacaaactctatttcggctttcgcggcggacgttt 1577
Db   1767 taatgatgataaacaagtagatccaaattcaatttactttggctttagaggcggtcgatt 1826

Qy   1578 ggatttaaacgggcattcgctttcgttccaccgtattcaaaataccgatgaaggggcgat 1637
Db   1827 agacttaaacggtaattcactaacctttgatcacatcagaaatattgatgatggtgcaag 1886

Qy   1638 gattgncnatcataatgccacaacaacatccaccgttaccattacagggaatgaaagtat 1697
```

```
Db   1887 actagttaatcataatatgactaatgcctcaaatataacgattactggggaaagtctaat 1946

Qy   1698 tacacaaccgagt 1710
          |||| | || | |
Db   1947 tacagatccaaat 1959
```

RESULT 24
ABA89172
ID ABA89172 standard; DNA; 4113 BP.
XX
AC ABA89172;
XX
DT 11-FEB-2002 (first entry)
XX
DE Escherichia coli polynucleotide SEQ ID NO 890.
XX
KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW systemic infection; non-diarrhoeal infection; septicaemia;
KW pyelonephritis; antibiotic resistance; ds.
XX
OS Escherichia coli.
XX
PN WO200166572-A2.
XX
PD 13-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-EP03445.
XX
PR 10-MAR-2000; 2000FR-0003145.
PR 02-FEB-2001; 2001FR-0001449.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
XX
DR WPI; 2001-550253/61.
XX
PT A library of DNA fragments of Escherichia coli strains for the
PT phylogenic determination of a given strain comprises polynucleotides of
PT nature B2/D+ A- -
XX
PS Example 6; Fig 6; 646pp; English.
XX
CC The invention relates to a library of DNA fragments of Escherichia coli
CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC B2/D+A-. The polynucleotides have potential antiinflammatory,
CC antibacterial and immunosuppressive activity as part of pharmaceutical
CC compositions used to treat, palliate or prevent extra-intestinal E. coli
CC infections. The polypeptides are useful for determining the phylogenic
CC group of a given E. coli strain. These polypeptides can detect and treat
CC an undesired development of E. coli, particularly an extra-intestinal
CC infection that include systemic and non-diarrhoeal infections such as
CC septicaemia, pyelonephritis and meningitis this is particularly
CC advantageous as bacterial resistance is increasing with the more
CC frequent use of broad spectrum antibiotics.
XX
SQ Sequence 4113 BP; 1112 A; 950 C; 1105 G; 946 T; 0 other;

```
  Query Match             3.2%;  Score 138.2;  DB 22;  Length 4113;
  Best Local Similarity   55.5%;  Pred. No. 3.1e-28;
  Matches  263;  Conservative    0;  Mismatches  211;  Indels    0;  Gaps    0;

Qy   1177 aacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttttgaa 1236
          |||||   | |  || |   | |||   | |  || || || ||   | |  | ||| ||
Db   1084 aacggtgcaattgtcctgaaagacagtgtgactcaggggccggttatctcgaatttaaa 1143

Qy   1237 ggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatc 1296
          |    ||  || || ||  |||||  |    |||| |||   || || ||| ||  ||
Db   1144 gacagttacaccgtatctgctgaatccggaaaaacctggacgggtgccggcattattact 1203

Qy   1297 agtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaa 1356
            | |  |  |  || || ||||| || ||||| || ||    |||  ||||   |||
Db   1204 gacaaggggacgaatgtgacctggaaggtcaacggggttgccggtgacaacctgcataaa 1263

Qy   1357 atcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagcgtg 1416
           | ||  |||| || |||  | |  | |  | ||| | |||||  ||     | |   |
Db   1264 ttgggggaaggaaccctgaccataaacggaacaggtgtaaacccgggaggactgaaaacg 1323

Qy   1417 ggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagccttt 1476
           || |||||||| ||  |  |  | ||||||||||||  |  ||| ||    || |||||
Db   1324 ggagacggtaccgttgtacttaaccagcaggcagacactgcaggtaatgttcaggccttc 1383

Qy   1477 agtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcagttc 1536
           |||    |  | |  ||||||  |    || |||  ||    |  |    |||| ||
Db   1384 agttccgtgaacctcgccagcggacgaccgaccgtggtgctcggagatgcccgtcaggtc 1443

Qy   1537 aaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcattcg 1596
           || || || ||  | |  |  || |  || || ||     | ||  | || ||  || |
Db   1444 aatccggataacatttcatgggggataccggggaggtaagcttgaccttaatggtaatgcc 1503

Qy   1597 ctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
           || | |||  |||  | ||   | ||||| | ||||||  |||||    || ||
Db   1504 gttaccttcacccgactgcaggctgccgattacggggcggtgattacaaataat 1557
```

RESULT 25
ABA89141/c
ID ABA89141 standard; DNA; 48254 BP.
XX
AC ABA89141;
XX
DT 11-FEB-2002 (first entry)
XX
DE Escherichia coli polynucleotide SEQ ID NO 829.
XX
KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW systemic infection; non-diarrhoeal infection; septicaemia;
KW pyelonephritis; antibiotic resistance; ds.
XX
OS Escherichia coli.
XX
PN WO200166572-A2.
XX
PD 13-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-EP03445.
XX
PR 10-MAR-2000; 2000FR-0003145.
PR 02-FEB-2001; 2001FR-0001449.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
XX
DR WPI; 2001-550253/61.
XX
PT A library of DNA fragments of Escherichia coli strains for the
PT phylogenic determination of a given strain comprises polynucleotides of
PT nature B2/D+ A- -
XX
PS Example 6; Fig 6; 646pp; English.
XX
CC The invention relates to a library of DNA fragments of Escherichia coli
CC strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC B2/D+A-. The polynucleotides have potential antiinflammatory,
CC antibacterial and immunosuppressive activity as part of pharmaceutical
CC compositions used to treat, palliate or prevent extra-intestinal E. coli
CC infections. The polypeptides are useful for determining the phylogenic
CC group of a given E. coli strain. These polypeptides can detect and treat
CC an undesired development of E. coli, particularly an extra-intestinal

```
CC   infection that include systemic and non-diarrhoeal infections such as
CC   septicaemia, pyelonephritis and meningitis this is particularly
CC   advantageous as bacterial resistance is increasing with the more
CC   frequent use of broad spectrum antibiotics.
XX
SQ   Sequence 48254 BP; 12342 A; 11284 C; 12417 G; 12211 T; 0 other;


  Query Match             3.2%;  Score 138.2;  DB 22;  Length 48254;
  Best Local Similarity   55.5%;  Pred. No. 1.3e-27;
  Matches  263;  Conservative    0;  Mismatches  211;  Indels    0;  Gaps    0;

Qy   1177 aacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttttgaa 1236
           |||||   | |  || |   | |||   | |  || || || ||   | |  | ||| ||
Db  28915 AACGGTGCAATTGTCCTGAAAGACAGTGTGACTCAGGGGGCCGGTTATCTCGAATTTAAA 28856

Qy   1237 ggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatc 1296
           |    ||  || || ||  |||||  |    |||| |||   || || ||| ||  ||
Db  28855 GACAGTTACACCGTATCTGCTGAATCCGGAAAAACCTGGACGGGTGCCGGCATTATTACT 28796

Qy   1297 agtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaa 1356
           | |  |  |  || || ||||| || ||||| || ||    |||  ||||   |||
Db  28795 GACAAGGGGACGAATGTGACCTGGAAGGTCAACGGGGTTGCCGGTGACAACCTGCATAAA 28736

Qy   1357 atcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagcgtg 1416
           | ||  |||| || |||  | |  | |  | ||| | |||||  ||     | |    |
Db  28735 TTGGGGGAAGGAACCCTGACCATAAACGGAACAGGTGTAAACCCGGGAGGACTGAAAACG 28676

Qy   1417 ggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagccttt 1476
           || |||||||| ||  |  |  | ||||||||||||  |  ||| ||    || |||||
Db  28675 GGAGACGGTACCGTTGTACTTAACCAGCAGGCAGACACTGCAGGTAATGTTCAGGCCTTC 28616

Qy   1477 agtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcagttc 1536
           |||    |  | |  ||||||  |    || |||   ||    |  |    |||| ||
Db  28615 AGTTCCGTGAACCTCGCCAGCGGACGACCGACCGTGGTGCTCGGAGATGCCCGTCAGGTC 28556

Qy   1537 aaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcattcg 1596
           || || || ||  | |  |  || |  || || ||     | ||  | || ||  || |
Db  28555 AATCCGGATAACATTTCATGGGGATACCGGGGAGGTAAGCTTGACCTTAATGGTAATGCC 28496

Qy   1597 ctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
           || | |||  |||  | ||   | ||||| | |||||| |||||    || ||
Db  28495 GTTACCTTCACCCGACTGCAGGCTGCCGATTACGGGGCGGTGATTACAAATAAT 28442


RESULT  26
ABA89142/c
ID   ABA89142 standard; DNA; 48345 BP.
XX
AC   ABA89142;
XX
DT   11-FEB-2002  (first entry)
XX
DE   Escherichia coli polynucleotide SEQ ID NO 830.
XX
KW   Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW   immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW   systemic infection; non-diarrhoeal infection; septicaemia;
KW   pyelonephritis; antibiotic resistance; ds.
XX
OS   Escherichia coli.
XX
PN   WO200166572-A2.
XX
PD   13-SEP-2001.
XX
PF   12-MAR-2001; 2001WO-EP03445.
XX
PR   10-MAR-2000; 2000FR-0003145.
PR   02-FEB-2001; 2001FR-0001449.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
XX
DR   WPI; 2001-550253/61.
XX
PT   A library of DNA fragments of Escherichia coli strains for the
PT   phylogenic determination of a given strain comprises polynucleotides of
PT   nature B2/D+ A- -
XX
PS   Example 6; Fig 6; 646pp; English.
XX
CC   The invention relates to a library of DNA fragments of Escherichia coli
CC   strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC   and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC   B2/D+A-. The polynucleotides have potential antiinflammatory,
CC   antibacterial and immunosuppressive activity as part of pharmaceutical
CC   compositions used to treat, palliate or prevent extra-intestinal E. coli
CC   infections. The polypeptides are useful for determining the phylogenic
CC   group of a given E. coli strain. These polypeptides can detect and treat
CC   an undesired development of E. coli, particularly an extra-intestinal
CC   infection that include systemic and non-diarrhoeal infections such as
CC   septicaemia, pyelonephritis and meningitis this is particularly
CC   advantageous as bacterial resistance is increasing with the more
CC   frequent use of broad spectrum antibiotics.
XX
SQ   Sequence 48345 BP; 12347 A; 11290 C; 12423 G; 12219 T; 66 other;


  Query Match             3.2%;  Score 138.2;  DB 22;  Length 48345;
  Best Local Similarity   55.5%;  Pred. No. 1.3e-27;
  Matches  263;  Conservative    0;  Mismatches  211;  Indels    0;  Gaps    0;

Qy   1177 aacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtatttttgaa 1236
           |||||   | |  || |   | |||   | |  || || || ||   | |  | ||| ||
Db  28915 AACGGTGCAATTGTCCTGAAAGACAGTGTGACTCAGGGGGCCGGTTATCTCGAATTTAAA 28856

Qy   1237 ggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatc 1296
           |    ||  || || ||  |||||  |    |||| |||   || || ||| ||  ||
Db  28855 GACAGTTACACCGTATCTGCTGAATCCGGAAAAACCTGGACGGGTGCCGGCATTATTACT 28796

Qy   1297 agtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaa 1356
           | |  |  |  || || ||||| || ||||| || ||    |||  ||||   |||
Db  28795 GACAAGGGGACGAATGTGACCTGGAAGGTCAACGGGGTTGCCGGTGACAACCTGCATAAA 28736

Qy   1357 atcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagcgtg 1416
           | ||  |||| || |||  | |  | |  | ||| | |||||  ||     | |    |
Db  28735 TTGGGGGAAGGAACCCTGACCATAAACGGAACAGGTGTAAACCCGGGAGGACTGAAAACG 28676

Qy   1417 ggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagccttt 1476
           || |||||||| ||  |  |  | ||||||||||||  |  ||| ||    || |||||
Db  28675 GGAGACGGTACCGTTGTACTTAACCAGCAGGCAGACACTGCAGGTAATGTTCAGGCCTTC 28616

Qy   1477 agtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcagttc 1536
           |||    |  | |  ||||||  |    || |||   ||    |  |    |||| ||
Db  28615 AGTTCCGTGAACCTCGCCAGCGGACGACCGACCGTGGTGCTCGGAGATGCCCGTCAGGTC 28556

Qy   1537 aaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcattcg 1596
           || || || ||  | |  |  || |  || || ||     | ||  | || ||  || |
Db  28555 AATCCGGATAACATTTCATGGGGATACCGGGGAGGTAAGCTTGACCTTAATGGTAATGCC 28496

Qy   1597 ctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
           || | |||  |||  | ||   | ||||| | |||||| |||||    || ||
Db  28495 GTTACCTTCACCCGACTGCAGGCTGCCGATTACGGGGCGGTGATTACAAATAAT 28442


RESULT  27
ABA88994
ID   ABA88994 standard; DNA; 4128 BP.
XX
AC   ABA88994;
XX
DT   11-FEB-2002  (first entry)
XX
DE   Escherichia coli polynucleotide SEQ ID NO 559.
```

```
XX
KW   Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW   immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW   systemic infection; non-diarrhoeal infection; septicaemia;
KW   pyelonephritis; antibiotic resistance; ds.
XX
OS   Escherichia coli.
XX
PN   WO200166572-A2.
XX
PD   13-SEP-2001.
XX
PF   12-MAR-2001; 2001WO-EP03445.
XX
PR   10-MAR-2000; 2000FR-0003145.
PR   02-FEB-2001; 2001FR-0001449.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Bingen E,  Bonacorsi S,  Clermont O,  Nassif X,  Tinsley C;
XX
DR   WPI; 2001-550253/61.
XX
PT   A library of DNA fragments of Escherichia coli strains for the
PT   phylogenic determination of a given strain comprises polynucleotides of
PT   nature B2/D+ A- -
XX
PS   Example 6; Fig 6; 646pp; English.
XX
CC   The invention relates to a library of DNA fragments of Escherichia coli
CC   strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC   and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC   B2/D+A-. The polynucleotides have potential antiinflammatory,
CC   antibacterial and immunosuppressive activity as part of pharmaceutical
CC   compositions used to treat, palliate or prevent extra-intestinal E. coli
CC   infections. The polypeptides are useful for determining the phylogenic
CC   group of a given E. coli strain. These polypeptides can detect and treat
CC   an undesired development of E. coli, particularly an extra-intestinal
CC   infection that include systemic and non-diarrhoeal infections such as
CC   septicaemia, pyelonephritis and meningitis this is particularly
CC   advantageous as bacterial resistance is increasing with the more
CC   frequent use of broad spectrum antibiotics.
XX
SQ   Sequence 4128 BP; 1151 A; 945 C; 1052 G; 980 T; 0 other;

  Query Match             2.8%;  Score 122;  DB 22;  Length 4128;
  Best Local Similarity   53.0%;  Pred. No. 1.3e-23;
  Matches  257;  Conservative    0;  Mismatches  228;  Indels    0;  Gaps    0;

Qy   1177 aacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtattttgaa 1236
           || ||  |  |   | |   ||||    |||  || || || ||     ||   |||
Db   1087 aatggtcagattgacctggaaaactctgtcacgcagggtgccggttcactgacatttact 1146

Qy   1237 ggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatc 1296
          | ||| |  || ||| |  ||  ||||      || |||   || || ||  ||  | |
Db   1147 gatgactacactgtcaccacttcaaacggaagtacctggaccggggccggtattattgtg 1206

Qy   1297 agtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaa 1356
               | ||     |||| |  ||| | || || || |||  |   |||  ||||   |||
Db   1207 gacaaggatgcctccgtaaactggcaggttaatggtgtgaaaggtgacaacctgcataaa 1266

Qy   1357 atcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagcgtg 1416
          |||||| |||| || |||   || || |  |  || |  ||  | |||    | |  ||
Db   1267 atcggcgaaggaaccctggttgtacagggaaccggtgttaatgagggcggcctgaaagtc 1326

Qy   1417 ggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagccttt 1476
          || || || || ||  |  |  |||||||||| |||  |  |||  |    || || ||
Db   1327 ggggatgggaccgttgtcctcaatcagcaggctgacagttcaggacacgttcaggcattc 1386

Qy   1477 agtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcagttc 1536
          |||    |     |  |||||||| |    || ||    |||   | | |  | ||| |
Db   1387 agtagcgtgaatattgccagcggccgcccgacagtcgtgctggcagacaaccagcaggtt 1446

Qy   1537 aaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcattcg 1596
          || || |||||  | |  |  ||||  || || ||   | |||||  | |||||| ||
Db   1447 aatccggacaatatatcctggggctaccgggggggggttctggatgttaacgggaatgac 1506

Qy   1597 ctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcataatgcc 1656
          ||  | || ||     |  |     ||||| | || || |   | |   |    | ||
Db   1507 ctgacatttcataagctgaatgccgccgattatggcgcaactctcggtaacagcagtgat 1566

Qy   1657 acaac 1661
          | |||
Db   1567 aaaac 1571


RESULT  28
ABA88991/c
ID   ABA88991 standard; DNA; 7654 BP.
XX
AC   ABA88991;
XX
DT   11-FEB-2002  (first entry)
XX
DE   Escherichia coli polynucleotide SEQ ID NO 554.
XX
KW   Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
KW   immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KW   systemic infection; non-diarrhoeal infection; septicaemia;
KW   pyelonephritis; antibiotic resistance; ds.
XX
OS   Escherichia coli.
XX
PN   WO200166572-A2.
XX
PD   13-SEP-2001.
XX
PF   12-MAR-2001; 2001WO-EP03445.
XX
PR   10-MAR-2000; 2000FR-0003145.
PR   02-FEB-2001; 2001FR-0001449.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Bingen E,  Bonacorsi S,  Clermont O,  Nassif X,  Tinsley C;
XX
DR   WPI; 2001-550253/61.
XX
PT   A library of DNA fragments of Escherichia coli strains for the
PT   phylogenic determination of a given strain comprises polynucleotides of
PT   nature B2/D+ A- -
XX
PS   Example 6; Fig 6; 646pp; English.
XX
CC   The invention relates to a library of DNA fragments of Escherichia coli
CC   strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533)
CC   and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC   B2/D+A-. The polynucleotides have potential antiinflammatory,
CC   antibacterial and immunosuppressive activity as part of pharmaceutical
CC   compositions used to treat, palliate or prevent extra-intestinal E. coli
CC   infections. The polypeptides are useful for determining the phylogenic
CC   group of a given E. coli strain. These polypeptides can detect and treat
CC   an undesired development of E. coli, particularly an extra-intestinal
CC   infection that include systemic and non-diarrhoeal infections such as
CC   septicaemia, pyelonephritis and meningitis this is particularly
CC   advantageous as bacterial resistance is increasing with the more
CC   frequent use of broad spectrum antibiotics.
XX
SQ   Sequence 7654 BP; 2000 A; 1697 C; 1656 G; 2301 T; 0 other;

  Query Match             2.8%;  Score 122;  DB 22;  Length 7654;
  Best Local Similarity   53.0%;  Pred. No. 1.8e-23;
  Matches  257;  Conservative    0;  Mismatches  228;  Indels    0;  Gaps    0;
```

```
Qy   1177 aacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtattttgaa 1236
          || ||  |  |   | |  ||||   |||  || || || ||     ||   |||
Db   4620 AATGGTCAGATTGACCTGGAAAACTCTGTCACGCAGGGTGCCGGTTCACTGACATTTACT 4561

Qy   1237 ggtgatttttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgttcatatc 1296
          | ||| |  || ||| |  ||  ||||      || |||   || || ||  ||  | |
Db   4560 GATGACTACACTGTCACCACTTCAAACGGAAGTACCTGGACCGGGGCCGGTATTATTGTG 4501

Qy   1297 agtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctgtccaaa 1356
               | ||     |||| |  ||| | || || || |||  |   |||  ||||   |||
Db   4500 GACAAGGATGCCTCCGTAAACTGGCAGGTTAATGGTGTGAAAGGTGACAACCTGCATAAA 4441

Qy   1357 atcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatcagcgtg 1416
          |||||| |||| || |||   || || |  |  || |  ||  | |||    | |  ||
Db   4440 ATCGGCGAAGGAACCCTGGTTGTACAGGGAACCGGTGTTAATGAGGGCGGCCTGAAAGTC 4381

Qy   1417 ggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaagccttt 1476
          || || || || ||  |  |  |||||||||| |||  |  |||  |    || || ||
Db   4380 GGGGATGGGACCGTTGTCCTCAATCAGCAGGCTGACAGTTCAGGACACGTTCAGGCATTC 4321

Qy   1477 agtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataatcagttc 1536
          |||    |     |   |||||||  |    || ||    |||   | | |  | ||| |
Db   4320 AGTAGCGTGAATATTGCCAGCGGCCGCCCGACAGTCGTGCTGGCAGACAACCAGCAGGTT 4261

Qy   1537 aaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacgggcattcg 1596
          || || |||||  | |  |  ||||  || || ||   | ||||| | |||||| ||
Db   4260 AATCCGGACAATATATCCTGGGGCTACCGGGGGGGGGTTCTGGATGTTAACGGGAATGAC 4201

Qy   1597 ctttcgttccaccgtattcaaaatacgatgaaggggcgatgattgncnatcataatgcc 1656
          ||  | || ||     |  |     ||||| | || || |   | |   |    | ||
Db   4200 CTGACATTTCATAAGCTGAATGCCGCCGATTATGGCGCAACTCTCGGTAACAGCAGTGAT 4141

Qy   1657 acaac 1661
          | |||
Db   4140 AAAAC 4136
```

RESULT 29
AAA81729
ID AAA81729 standard; DNA; 642 BP.
XX
AC AAA81729;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_276 SEQ ID NO:276.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
PS Claim 7; Page 1554-1555; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neissariae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
SQ Sequence 642 BP; 131 A; 145 C; 193 G; 173 T; 0 other;

```
  Query Match             2.5%;  Score 108.2;  DB 21;  Length 642;
  Best Local Similarity   77.8%;  Pred. No. 3.6e-20;
  Matches  144;  Conservative    0;  Mismatches   35;  Indels    6;  Gaps    1;

Qy   1587 cgggcattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncna 1646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db    464 cgggcattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgtcaa 523

Qy   1647 tcataatgccacaacaacatccaccgttaccattacagggaatgaaagtattacacaacc 1706
            || |||      | |  |||||||||||||||||||||| ||| ||  |||| |   | |
Db    524 ccacaatcaagacaaagaatccaccgttaccattacaggcaataaagatattgctacaac 583

Qy   1707 gagtggtaagaatatcaatagacttaattacagcaaagaaattgcctacaacggttggtt 1766
            |       |||| ||| ||  |  ||  ||  |||||||||||||||||||||||||
Db    584 cgg------caataacaacagcttggatagcaaaaaagaaattgcctacaacggttggtt 637

Qy   1767 tggcg 1771
          |||||
Db    638 tggcg 642
```

RESULT 30
AAA10594
ID AAA10594 standard; DNA; 10732 BP.
XX
AC AAA10594;
XX
DT 29-JUN-2000 (first entry)
XX
DE Gene encoding a subunit of cellulose synthase.
XX
KW Cellulose synthase; cellulose production; increase yield; ds.
XX
OS Vigna angularis.
XX
PN JP2000060568-A.
XX
PD 29-FEB-2000.
XX
PF 26-AUG-1998; 98JP-0239998.
XX

PR 26-AUG-1998; 98JP-0239998.
XX
PA (MIZU/) MIZUNO K.
PA (OJIP) OJI PAPER CO.
XX
DR WPI; 2000-342371/30.
DR P-PSDB; AAY85179.
XX
PT A gene encoding a cellulose synthetic equipment - for the improvement
PT in the amount of cellulose synthesised in a plant body
XX
PS Claim 2; Page 14-21; 32pp; Japanese.
XX
CC This sequence represents a gene encoding a subunit of the cellulose
CC synthase complex of Vigna angularis. The invention relates to subunits of
CC cellulose synthetic equipment, that can be used to increase the amount of
CC cellulose synthesised by a plant. The proteins and genes encoding them
CC can also be used to improve the properties of the cellulose being
CC produced by a plant.
XX
SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;

Query Match 1.2%; Score 52.2; DB 21; Length 10732;
Best Local Similarity 13.7%; Pred. No. 0.0017;
Matches 128; Conservative 367; Mismatches 428; Indels 14; Gaps 2;

Qy 3067 aacgaacacgtcgatgccggcgcgtggcgttaccaactcatccgcaaagacggcgagttc 3126
Db 9276 argthraagvasrhsysgngngsrthrysysaaysggcysgnargargvaargysargrg 9335

Qy 3127 cgcctgcataatccggtcaaagaacaagagctttccgacaaactcggcaaggcagaagcc 3186
Db 9336 yraaaaaaysmtysasngvagmtgyargassrhgargargsryssrthrasnsrvagsrs 9395

Qy 3187 aaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctgattgcggccgggcgc 3246
Db 9396 rvaasthrsrrgthrrthrysasnthrthrggncysaavagggasnysvaysgsrasnar 9455

Qy 3247 gatgccgccgaaaagacagaaagcgttgccgaaccggcccggcnggcaggcggggaaaat 3306
Db 9456 gysasnggnhsrargvamtthrargthraasryshsgsrgngashsysgyhssrvaghsh 9515

Qy 3307 gtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcggataaagacagcgcn 3366
Db 9516 ssrasnvaaasrhsghsrhaasrsrasaagysrasasysthrsrcysaagsrtraasraa 9575

Qy 3367 ttggcgaaacagcgcgaagcggaaacccggccggntaccaccgccttcccccgcgcccgc 3426
Db 9576 srgashshargsrgysysysgrsrcysyssrvagyaasrasasnmtasashvagmtgysa 9635

Qy 3427 ngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacctcaaccccaaccgcag 3486
Db 9636 avavasrvagysaathrgnsrrsraasrysaavathrgasnsrsrgthrgythrysgsrt 9695

Qy 3487 cgcgacctgatnagccgttatgccaatagcggtttgagtgaattttccgccacgctcaac 3546
Db 9696 hrrgggyysgrvarashssrgnthrthrgthrthrrgvavaasmtgrmtsrashssrass 9755

Qy 3547 agcgtt----------ttcgccgtacaggacgaattggaccgcgtgtttgccgaaga 3593
Db 9756 ryssrysasnthrcyssrashysgythrvargytrgnasvavaysthrggnasnhsvath 9815

Qy 3594 ccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaaga 3653
Db 9816 rhsysgyrasashsasysvaaaarghsvaasnashashashassrsrarggysrvahsas 9875

Qy 3654 tttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctcgg 3713
Db 9876 thrgnasnrrgncyshscyssrcyssrassrsrvavaasnrsraagygasnaaassrysa 9935

Qy 3714 cagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgacgg 3773
Db 9936 rggrsrgngassryssrgyysgvagargraavaastyrassrsrasrgnysgygyargas 9995

Qy 3774 catcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggcag 3833
Db 9996 hsrtyrysasngymtrthrgytyrmtvaargvahgntrysthrsrgsrysvagghhsvac 10055

Qy 3834 gttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggcat 3893
Db 10056 ystyrassrgyysthrastyrgyasnhaagythrthraaastrmtasnhscyshsrgnas 10115

Qy 3894 cggaggcaaaatccgccgccgcgtgctgcattacggcattcaggcacgataccgcgccgg 3953
Db 10116 vasrsrm-targasaaysysgnhastrasgthrargsrasgyggasngthrhsvagysga 10174

Qy 3954 tttcggcggattcggcatcgaaccgtacatcggcgca 3990
Db 10175 rghsrarggasnsrsrrgnvathrthrsrcysgyhsa 10211

RESULT 31
AAA30290
ID AAA30290 standard; DNA; 3489 BP.
XX
AC AAA30290;
XX
DT 11-SEP-2000 (first entry)
XX
DE Kaposi's sarcoma-associated herpesvirus LANA gene.
XX
KW Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
KW latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
KW Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
KW Kaposi's sarcoma; primary effusion lymphoma; PEL;
KW human immunodeficiency virus; HIV; multicentric Castleman's disease; ds.
XX
OS Kaposi's sarcoma-associated herpesvirus.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /*tag= a
FT /product= "LANA"
FT misc_signal 40..50
FT /*tag= b
FT /note= "nuclear localisation signal, NLS"
FT misc_signal 190..210
FT /*tag= c
FT /note= "nuclear localisation signal, NLS"
XX
PN WO200029626-A1.
XX
PD 25-MAY-2000.
XX
PF 19-NOV-1999; 99WO-US27508.
XX
PR 19-NOV-1998; 98US-0109422.
PR 21-APR-1999; 99US-0298568.
XX
PA (KIEF/) KIEFF E D.
PA (BALL/) BALLESTAS M E.
PA (KAYE/) KAYE K M.
XX
PI Kieff ED, Ballestas ME, Kaye KM;
XX
DR WPI; 2000-387829/33.
DR P-PSDB; AAY96255.
XX
PT Treating or preventing a disease associated with rhodino virus
PT infection in a mammal which includes Kaposi's Sarcoma and Primary
PT Effusion Lymphoma -
XX
PS Disclosure; Fig 6; 70pp; English.
XX
CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,

CC (KSHV) latency-associated nuclear antigen (LANA) gene. KSHV is also known
CC as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or
CC gamma-2 herpes virus class. The LANA protein is necessary for the
CC efficient persistence of rhadino virus DNA in mammalian cells. Persistent
CC rhadino virus infection is implicated in a variety of diseases e.g.
CC Kaposi's Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
CC Castleman's disease. In addition, KS is a common malignancy in HIV
CC patients. KSHV persists in host cells in a latent form. One of the few
CC genes expressed from the latent viral DNA is LANA. LANA associates with
CC both human chromosomes and with the rhadino virus cis-acting element
CC (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
CC "tied" to the host chromosomes. This allows the viral DNA to persist in
CC the host cell. The present sequence may be used to screen and identify
CC molecules that inhibit LANA interaction with RVCAE, thereby interfering
CC with the latency cycle of this virus. Potential antiviral treatments for
CC the above mentioned diseases may therefore be based on LANA deregulation.
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

Query Match 1.1%; Score 47.6; DB 21; Length 3489;
Best Local Similarity 42.4%; Pred. No. 0.017;
Matches 342; Conservative 0; Mismatches 460; Indels 5; Gaps 2;

```
Qy   3112 aaagacggcgagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactc 3171
          | ||| |  |||  |     | || |   ||   | ||  | |||  |  |||  |
Db   1009 acagatgaggaggacgaggaggatgacgaggaggatgacgaggaggatgacgaggaggat 1068

Qy   3172 ggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231
          | | |||  || | | |  |  | |  ||  | ||  ||| | |      |||| |  |
Db   1069 gacgaggaggatgacgaggaggatgacgaggaggatgacgagga---ggatgacgaggag 1125

Qy   3232 attgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcng 3291
            ||  |  | |   || |  |  ||  |    ||    |  |  ||   || |  |  |
Db   1126 gatgacgaggaggatgacgaggaggatgacgaggaggaggacgaggaggaggacgaggag 1185

Qy   3292 gcaggcgggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3351
          |  | || |||  | |  | |     | | |  || || ||| | ||   ||   ||| |
Db   1186 gaggacgaggaggaggaggacgaggaggatgacgatgatgaggacaatgaggacgaggag 1245

Qy   3352 gataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgcc 3411
          ||| | ||    |       |||  ||  ||| | |||     |    |  | ||  |
Db   1246 gatgacgaggaggaggacaagaaggaggacgaggaggacgggggcgatggaaacaaaacg 1305

Qy   3412 ttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacct 3471
          ||   |  |    |       |  | | |   ||| |||| ||||  |  || ||| |
Db   1306 ttgagcatccaaagttcacaacagcagcaggagccacaacagcaggagccacagcagcag 1365

Qy   3472 caaccccaaccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattt 3531
            | || || | ||||    ||||  ||  |  |  ||  ||  |    ||
Db   1366 gagccacagcagcaggagcccctg--caggagccacaacagcaggagccacagcagcagg 1423

Qy   3532 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3591
            |  ||    ||  |||   | |      ||  ||     |  || |
Db   1424 agccacagcagcaggagcccctgcaggagccacaacagcaggagccacagcagcaggagc 1483

Qy   3592 gaccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaa 3651
            | || | |       | |  |||  || | |    | | |   |  | |    ||
Db   1484 ccctgcaggagccacaacagcaggagccacaacagcaggagccacagcagcaggagccac 1543

Qy   3652 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3711
             | |   | || ||      | | ||    ||   | |   |  |  | | || |
Db   1544 agcagcaggagccacagcagcaggagccacagcagcaggagccacagcagcaggagccac 1603

Qy   3712 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgac 3771
          |||||  || ||  | ||| |  |        || | ||| | | | || |     | ||
Db   1604 agcagcaggagccacagcagcaggagccacagcagcaggagccacagcagcgggagccac 1663

Qy   3772 ggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggc 3831
          ||| ||| | | |      |      |||| ||  ||    |        ||| | |  |
Db   1664 agcagcgggagccccagcagcgggagccacagcagcgggagccacagcagcgggagccac 1723

Qy   3832 aggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggc 3891
```

```
          ||  ||  | |  || ||||  | ||    |    | |||  ||         |  |
Db   1724 agcagcgggagccacagcagcgggagccacagcagcgggagccacagcagcgggagccac 1783

Qy   3892 atcggaggcaaaatccgccgccgcgtg 3918
          | | |  | |  | | || || |  ||
Db   1784 agcagcaggatgagcagcagcaggatg 1810
```

RESULT 32
AAF82901
ID AAF82901 standard; DNA; 3489 BP.
XX
AC AAF82901;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS Kaposi's sarcoma associated herpesvirus.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /*tag= a
XX
PN WO200125484-A2.
XX
PD 12-APR-2001.
XX
PF 29-SEP-2000; 2000WO-US26908.
XX
PR 01-OCT-1999; 99US-0410399.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Robertson ES, Cotter MA;
XX
DR WPI; 2001-281736/29.
DR P-PSDB; AAB62331.
XX
PT A composition for use in gene therapy comprises an expression vector
PT that includes a nucleic acid sequence encoding a nucleic acid binding
PT protein -
XX
PS Disclosure; Fig 9A; 60pp; English.
XX
CC The invention provides a composition comprising nucleic acid, histone H1
CC protein and expression vector operationally encoding a protein suitable
CC for tethering the nucleic acid to the histone H1 protein, where the
CC tethering protein is LANA. The composition is useful in aiding the
CC retention of the viral DNA in the host cell. The viral vector encodes a
CC protein suitable for tethering DNA to Histone H1. Methods for screening
CC for compounds which are agonistic or antagonistic for the tethering of
CC viral proteins to histone H1 and DNA binding sites are useful for
CC developing the method of viral transfer. The composition has applications
CC to gene therapy, including the treatment of multiple sclerosis,
CC Parkinson's disease, Huntington disease and diabetes. The present
CC sequence represents the nucleotide sequence of the Kaposi's sarcoma
CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC antigen (LANA), which acts as a tethering protein.
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

Query Match 1.1%; Score 47.6; DB 22; Length 3489;
Best Local Similarity 42.4%; Pred. No. 0.017;
Matches 342; Conservative 0; Mismatches 460; Indels 5; Gaps 2;

```
Qy   3112 aaagacggcgagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactc 3171
          | ||| |  |||  |     | || |   ||   | ||  | |||  |  |||  |
```

```
Db   1009 acagatgaggaggacgaggaggatgacgaggaggatgacgaggaggatgacgaggaggat 1068

Qy   3172 ggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231

Db   1069 gacgaggaggatgacgaggaggatgacgaggaggatgacgagga---ggatgacgaggag 1125

Qy   3232 attgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcng 3291

Db   1126 gatgacgaggaggatgacgaggaggatgacgaggaggaggacgaggaggaggacgaggag 1185

Qy   3292 gcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3351

Db   1186 gaggacgaggaggaggaggacgaggaggatgacgatgatgaggacaatgaggacgaggag 1245

Qy   3352 gataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgcc 3411

Db   1246 gatgacgaggaggaggacaagaaggaggacgaggaggacgggggcgatggaaacaaaacg 1305

Qy   3412 ttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacct 3471

Db   1306 ttgagcatccaaagttcacaacagcagcaggagccacaacagcaggagccacagcagcag 1365

Qy   3472 caaccccaaccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattt 3531

Db   1366 gagccacagcagcaggagcccctg--caggagccacaacagcaggagccacagcagcagg 1423

Qy   3532 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3591

Db   1424 agccacagcagcaggagcccctgcaggagccacaacagcaggagccacagcagcaggagc 1483

Qy   3592 gaccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaa 3651

Db   1484 ccctgcaggagccacaacagcaggagccacaacagcaggagccacagcagcaggagccac 1543

Qy   3652 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3711

Db   1544 agcagcaggagccacagcagcaggagccacagcagcaggagccacagcagcaggagccac 1603

Qy   3712 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgac 3771

Db   1604 agcagcaggagccacagcagcaggagccacagcagcaggagccacagcagcgggagccac 1663

Qy   3772 ggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggc 3831

Db   1664 agcagcgggagccccagcagcgggagccacagcagcgggagccacagcagcgggagccac 1723

Qy   3832 aggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggc 3891

Db   1724 agcagcgggagccacagcagcgggagccacagcagcgggagccacagcagcgggagccac 1783

Qy   3892 atcggaggcaaaatccgccgccgcgtg 3918

Db   1784 agcagcaggatgagcagcagcaggatg 1810
```

```
RESULT 33
AAV73805/c
ID   AAV73805 standard; DNA; 32207 BP.
XX
AC   AAV73805;
XX
DT   25-FEB-1999  (first entry)
XX
DE   KSHV LUR DNA (nucleotides 105,301-137,507).
XX
KW   Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;
KW   dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;
KW   diagnosis; treatment; HHV8; capsid protein IV; tegument protein IV;
KW   glycoprotein; kaposin; cyclin D; immediate early protein; IEP; OX-2;
KW   v-adh; G-protein coupled receptor; FGARAT; ds.
XX
OS   Kaposi's sarcoma-associated herpesvirus.
XX
PN   US5849564-A.
XX
PD   15-DEC-1998.
XX
PF   29-NOV-1996;   96US-0770379.
XX
PR   29-NOV-1996;   96US-0770379.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Bohenzky RA,  Chang Y,  Edelman IS,  Moore PS,  Russo JJ;
XX
DR   WPI; 1999-069741/06.
XX
PT   Kaposi's sarcoma-associated herpes virus nucleic acid - encodes
PT   di:hydro:folate reductase and is useful for treatment, prophylaxis
PT   or diagnosis of Kaposi's sarcoma
XX
PS   Disclosure; Column 155-182; 109pp; English.
XX
CC   This sequence is a fragment of the Kaposi's sarcoma-associated
CC   herpesvirus (KSHV) LUR (long unique region). This fragment contains
CC   coding regions for ORF65 which encodes capsid protein IV, ORF66, ORF67
CC   which encodes tegument protein IV, ORF68 which encodes a glycoprotein,
CC   ORF69, K12 which encodes kaposin, K13, ORF72 which encodes cyclin D,
CC   ORF73 which encodes immediate early protein (IEP), K14 which encodes
CC   OX-2 (v-adh), ORF74 which encodes G-protein coupled receptor, ORF75
CC   which encodes tegument protein/FGARAT, K15. KSHV is a new human
CC   Herpesvirus (HHV8) believed to cause Kaposi's sarcoma (KS) which is the
CC   most common form of neoplasm occurring in persons with acquired immune
CC   deficiency syndrome (AIDS). The DHFR protein is useful for vaccination,
CC   prophylaxis, diagnosis and treatment of a subject with Kaposi's sarcoma
CC   and for detecting expression of a DNA virus associated with Kaposi's
CC   sarcoma in a cell.
XX
SQ   Sequence 32207 BP; 7229 A; 9156 C; 8713 G; 7109 T; 0 other;
```

```
  Query Match             1.1%;  Score 47.6;  DB 20;  Length 32207;
  Best Local Similarity   42.4%;  Pred. No. 0.064;
  Matches  342;  Conservative    0;  Mismatches  460;  Indels    5;  Gaps    2;

Qy   3112 aaagacggcgagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactc 3171

Db  20988 ACAGATGAGGAGGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAT 20929

Qy   3172 ggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231

Db  20928 GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGA---GGATGACGAGGAG 20872

Qy   3232 attgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcng 3291

Db  20871 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAGGACGAGGAGGAGGACGAGGAG 20812

Qy   3292 gcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3351

Db  20811 GAGGACGAGGAGGAGGAGGACGAGGAGGATGACGATGATGAGGACAATGAGGACGAGGAG 20752

Qy   3352 gataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgcc 3411

Db  20751 GATGACGAGGAGGAGGACAAGAAGGAGGACGAGGAGGACGGGGGCGATGGAAACAAAACG 20692

Qy   3412 ttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacct 3471

Db  20691 TTGAGCATCCAAAGTTCACAACAGCAGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAG 20632

Qy   3472 caaccccaaccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattt 3531

Db  20631 GAGCCACAGCAGCAGGAGCCCCTG--CAGGAGCCACAACAGCAGGAGCCACAGCAGCAGG 20574

Qy   3532 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3591

Db  20573 AGCCACAGCAGCAGGAGCCCCTGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAGGAGC 20514
```

```
Qy    3592 gaccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaa 3651
           | || | | |       | |  |||  || | |    | | |   |  | |    ||
Db   20513 CCCTGCAGGAGCCACAACAGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAGGAGCCAC 20454

Qy    3652 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3711
           | |   | || ||      | | ||    ||   | |   |     |  | || |
Db   20453 AGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCAC 20394

Qy    3712 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgac 3771
           ||||| || ||  | ||| |  |         || | ||| | | | || |     | ||
Db   20393 AGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCGGGAGCCAC 20334

Qy    3772 ggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggc 3831
           ||| ||| | | |     |      |||| ||  ||    |       ||| | |  |
Db   20333 AGCAGCGGGAGCCCCAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCAC 20274

Qy    3832 aggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggc 3891
           ||   ||  | |  || |||| | ||    |     | |||  ||         |  |
Db   20273 AGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCAC 20214

Qy    3892 atcggaggcaaaatccgccgccgcgtg 3918
           | | |  | |  | | || || |  ||
Db   20213 AGCAGCAGGATGAGCAGCAGCAGGATG 20187
```

RESULT 34
AAV19941/c
ID AAV19941 standard; DNA; 137507 BP.
XX
AC AAV19941;
XX
DT 03-AUG-1998 (first entry)
XX
DE KSHV long unique coding region and terminal repeat.
XX
KW KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;
KW interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;
KW complement-binding protein; glycoprotein; capsid protein IV; infection;
KW immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;
KW lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;
KW HIV immune status; anti-inflammatory agent; therapy; ds.
XX
OS Kaposi's sarcoma-associated herpes virus.
XX

FH	Key	Location/Qualifiers
FT	CDS	1142..2794
FT		/*tag= a
FT		/product= complement-binding protein
FT	CDS	8699..11236
FT		/*tag= b
FT		/product= glycoprotein B
FT	CDS	complement (17261..17875)
FT		/*tag= c
FT		/product= interleukin 6
FT	CDS	complement (21548..21832)
FT		/*tag= d
FT		/product= macrophage inflammatory protein II
FT	CDS	complement (27137..27424)
FT		/*tag= e
FT		/product= interferon regulatory factor 1
FT	CDS	28661..29741
FT		/*tag= f
FT		/product= protein T1.1
FT	CDS	complement (58976..60175)
FT		/*tag= g
FT		/product= glycoprotein M
FT	CDS	complement (69412..69915)
FT		/*tag= h
FT		/product= glycoprotein L
FT	CDS	complement (88410..88910)
FT		/*tag= i
FT		/product= interferon regulatory factor 2
FT	CDS	89600..90541
FT		/*tag= j
FT		/product= interferon regulatory factor 3
FT	CDS	90173..90643
FT		/*tag= k
FT		/product= glycoprotein X
FT	CDS	complement (93636..94127)
FT		/*tag= l
FT		/product= interferon regulatory factor 4
FT	CDS	complement (111931..112443)
FT		/*tag= m
FT		/product= capsid protein IV
FT	CDS	complement (123808..127296)
FT		/*tag= n
FT		/product= immediate early protein

XX
PN WO9804576-A1.
XX
PD 05-FEB-1998.
XX
PF 22-JUL-1997; 97WO-US13346.
XX
PR 29-NOV-1996; 96US-0757669.
PR 25-JUL-1996; 96US-0686243.
PR 25-JUL-1996; 96US-0686349.
PR 25-JUL-1996; 96US-0686350.
PR 25-JUL-1996; 96US-0687253.
PR 25-JUL-1996; 96US-0688814.
PR 05-SEP-1996; 96US-0708678.
PR 10-OCT-1996; 96US-0728323.
PR 13-NOV-1996; 96US-0747887.
PR 13-NOV-1996; 96US-0748640.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Bohenzky RA, Chang Y, Edelman IS, Moore PS, Russo JJ;
XX
DR WPI; 1998-130615/12.
XX
PT New nucleic acid encoding Kaposi's sarcoma associated herpes virus
PT proteins - useful for, e.g. detecting levels of HHV8 in, and
PT preparation of vaccines for treatment of, HIV patients
XX
PS Example 2; Page 135-203; 230pp; English.
XX
CC This sequence represents the long unique region and terminal repeat of
CC the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
CC as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
CC invention which encode KSHV polypeptides selected from: (a) viral
CC macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
CC (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
CC (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
CC encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
CC by it, and antibodies (Ab) specific for the proteins are useful for
CC detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
CC fluids or tissue samples. HHV8 infections can be treated with antisense
CC or triplex forming molecules or agents that bind specifically to the
CC protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
CC while the protein can be used in protective vaccines. Ab may also be used
CC to differentiate between lymphomas, and HHV8 may be implicated in many
CC other lymphoproliferative diseases such as lymphomas, leukaemia,
CC splenomegaly and mycosis fungoides. Cells and animals containing the
CC nucleic acid are useful for drug screening. HHV8-derived peptides can be
CC used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
CC can be inhibited with methotrexate. These can also be used to determine
CC the immune status of a patient infected with HIV. HHV8 derived protein
CC viral MIP III may be used as an anti-inflammatory agent for,
CC e.g. treating rheumatoid arthritis. This sequence is stated as containing
CC 81 open reading frames.
XX
SQ Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 other;

```
  Query Match             1.1%;  Score 47.6;  DB 19;  Length 137507;
  Best Local Similarity   42.4%;  Pred. No. 0.15;
  Matches  342;  Conservative    0;  Mismatches  460;  Indels    5;  Gaps    2;

Qy   3112 aaagacggcgagttccgcctgcataatccggtcaaagaacaagagctttccgacaaactc 3171
          | ||| |  |||  |     | || |   ||   | ||  | |||  |  |||  |
Db 126288 ACAGATGAGGAGGACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAT 126229

Qy   3172 ggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaaagccttgacgcgctg 3231
          | | |||  || | | |  |  | | |  ||  | ||  ||| | |      ||||| |  |
Db 126228 GACGAGGAGGATGACGAGGAGGATGACGAGGAGGATGACGAGGA---GGATGACGAGGAG 126172

Qy   3232 attgcggccgggcgcgatgccgccgaaaagacagaaagcgttgccgaaccggcccggcng 3291
            ||  |  | |   || | |  ||  |    ||    |  |  ||   || | |  |
Db 126171 GATGACGAGGAGGATGACGAGGAGGATGACGAGGAGGAGGACGAGGAGGAGGACGAGGAG 126112

Qy   3292 gcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaaaaacgggtgcaggcg 3351
          |  | || |||  | |  | |     | | |  || || ||| | ||   ||   ||| |
Db 126111 GAGGACGAGGAGGAGGAGGACGAGGAGGATGACGATGATGAGGACAATGAGGACGAGGAG 126052

Qy   3352 gataaagacagcgcnttggcgaaacagcgcgaagcggaaacccggccggntaccaccgcc 3411
          ||| | ||    |       |||  ||  ||| | |||      |    |  | ||   |
Db 126051 GATGACGAGGAGGAGGACAAGAAGGAGGACGAGGAGGACGGGGGCGATGGAAACAAAACG 125992

Qy   3412 ttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgcagccccaaccgcaacct 3471
          ||   |  |    |       |  | | |   ||| |||| ||||   | || ||| |
Db 125991 TTGAGCATCCAAAGTTCACAACAGCAGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAG 125932

Qy   3472 caaccccaaccgcagcgcgacctgatnagccgttatgccaatagcggtttgagtgaattt 3531
            | || || | ||||     ||||   ||  |  |   ||  ||  |    ||
Db 125931 GAGCCACAGCAGCAGGAGCCCCTG--CAGGAGCCACAACAGCAGGAGCCACAGCAGCAGG 125874

Qy   3532 tccgccacgctcaacagcgttttcgccgtacaggacgaattggaccgcgtgtttgccgaa 3591
            |  ||    ||  |||   | |      ||  ||      |  || |
Db 125873 AGCCACAGCAGCAGGAGCCCCTGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAGGAGC 125814

Qy   3592 gaccgccgcaacgcngtttggacaagcngcatccggnacaccaaacactaccgttcgcaa 3651
            | || | | |       | |  |||  || | |    | | |   |  | |    ||
Db 125813 CCCTGCAGGAGCCACAACAGCAGGAGCCACAACAGCAGGAGCCACAGCAGCAGGAGCCAC 125754

Qy   3652 gatttccgcgcctaccgccaacaaaccgacctgcgccaaatcggtatgcagaaaaacctc 3711
              | |   | || ||      | | ||    ||   | |   |     |  | || |
Db 125753 AGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCAC 125694

Qy   3712 ggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttcgacgac 3771
          |||||| || ||  | ||| |  |        || | ||| | | | || |     | ||
Db 125693 AGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCGGGAGCCAC 125634

Qy   3772 ggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggcatcggc 3831
          ||| ||| | | |     |      |||| ||  ||      |        ||| | |  |
Db 125633 AGCAGCGGGAGCCCCAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCAC 125574

Qy   3832 aggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntcagacggc 3891
          ||   ||  | |  || ||||  | ||    |     | |||  ||          |  |
Db 125573 AGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCAC 125514

Qy   3892 atcggaggcaaaatccgccgccgcgtg 3918
          | | |  | |  | | || || |  ||
Db 125513 AGCAGCAGGATGAGCAGCAGCAGGATG 125487


RESULT  35
AAI99683
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;


  Query Match             1.0%;  Score 43.8;  DB 22;  Length 4403765;
  Best Local Similarity   43.0%;  Pred. No. 14;
  Matches  205;  Conservative    0;  Mismatches  272;  Indels    0;  Gaps    0;

Qy   3576 ccgcgtgtttgccgaagaccgccgcaacgcngtttggacaagcngcatccggnacaccaa 3635
          | |||     ||||  | | || ||| ||   |       | ||  |  | |  || ||
Db 3934543 cggcggcaacgccggcgccggcggcaccggcttcacccaaggcgccgacggcaacgccgg 3934602

Qy   3636 acactaccgttcgcaagatttccgcgcctaccgccaacaaaccgacctgcgccaaatcgg 3695
            ||  | ||      |   || ||| | || ||  | |   ||       |   | |
Db 3934603 caacggcggtgacggcggggtcggcggcaacggcggaaacggcgcagacaacaccaccac 3934662

Qy   3696 tatgcagaaaaacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccga 3755
                      |  |    |||||  || |||   |       |   |       ||
Db 3934663 cgccgccgccggcaccacaggcggggccggcggggccggcggggccggcggaaccggcgg 3934722

Qy   3756 aaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgttttcgg 3815
          || |  |  || || ||||| ||||| |   ||  | |  | | ||||||  |     |||
Db 3934723 aaccggcggcgccgccggcaccggcagcggcgccggacaaggcaacggcggcaacggcgg 3934782

Qy   3816 gcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcagcgg 3875
            |   ||||| |||| |    | |  |  |  | |  | |||| |       |  |
Db 3934783 caacggcggcaccggcggcaaaggcggcaccggcggggccggcgtgaacaacctcgacca 3934842

Qy   3876 cantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcattca 3935
            |  | | |  | ||| ||| ||  | ||  | |||          |  |
Db 3934843 ggtggtagccgcccacgacggcggccaaggcggcaccggcggcaccggcggcaacgccgg 3934902

Qy   3936 ggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaacgcg 3995
            ||  |   |  || |     |   ||   || |  | |  |   || |||||      |
Db 3934903 cgccggcggcaccggcttcacccaaggcgccgacggcaacgccggcaacggcggtgacgg 3934962
```

```
Qy   3996 ctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccgg 4052
          |    ||| |   || |  |   || ||   || |||  ||  | |||| || ||||
Db 3934963 cggggtcggcggcaacggcggaaacggcgcagacaacaccaccaccgccgccgccgg 3935019


RESULT 36
AAI99682
ID   AAI99682 standard; DNA; 4411529 BP.
XX
AC   AAI99682;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;


  Query Match             1.0%;  Score 43.8;  DB 22;  Length 4411529;
  Best Local Similarity   46.3%;  Pred. No. 14;
  Matches  174;  Conservative    0;  Mismatches  199;  Indels    3;  Gaps    1;

Qy   3706 aacctcggcagcgggcgcgtcggcatcctgttttcgcacaaccggaccgaaaacancttc 3765
          | |  ||||| |||| |||| |||  |          ||  | | | || |  || |  |
Db 3949060 accggcggcaccgggggcgtgggcggcaccgggggcgacggcggcaacgcaggcaccggc 3949119

Qy   3766 gacgacggcatcggcaactcggcacggcttgcccacggcgccgttttcgggcaatacggc 3825
          | || || |  ||||||     |       | |  ||||| |     |||      |||
Db 3949120 gccggcgacccccggcaaaggcggcaccggcggcaccggcggcaccggcggcagcggcggg 3949179

Qy   3826 atcggcaggttcgacatcggcatcagcacgggcgcgggttttagcagcggcantctntca 3885
          ||||  |   || |   | || |  ||  ||||         ||| ||||       |
Db 3949180 gccggcggtagcggcggggccaacttcaacggcggcaccggcggcaccggcgg---cacc 3949236
```

```
Qy   3886 gacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcattcaggcacgatac 3945
          | ||||| ||| ||||||  | ||    |||   |   |  || |     | |  |
Db 3949237 ggcggcaccggcggcaaaggcggcatgggcggcatcgctggcgacggcgggcccggcggt 3949296

Qy   3946 cgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaacgcgctatttcgtc 4005
            || |||   || |||  ||||     ||     || |  ||  |   ||     || |
Db 3949297 gacggcggcaacgccggggtcggaggaaaaggcggcaccaacggcaacggcggcagcggc 3949356

Qy   4006 caaaaagcggattaccgctacgaaaacgtcaatatcgccacccccggtcttgcgttcaac 4065
            |  |  |       ||   |    || |||   || | ||  ||| || ||   || |
Db 3949357 gggaccggcggcacaggcggggccggcggcaacgccggcgccggcggcctagccaacacc 3949416

Qy   4066 cgntaccgngcgggca 4081
          |   |   || ||||
Db 3949417 ggcggcaccgcaggca 3949432


RESULT 37
AAH98282
ID   AAH98282 standard; cDNA; 2068 BP.
XX
AC   AAH98282;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Rabbit EST-derived coding sequence SEQ ID NO: 139.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition; ss.
XX
OS   Oryctolagus cuniculus.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   P-PSDB; AAM23623.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 1; Page 265-266; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a cDNA
CC   of the invention.
XX
SQ   Sequence 2068 BP; 465 A; 521 C; 829 G; 253 T; 0 other;


  Query Match             1.0%;  Score 43.6;  DB 22;  Length 2068;
  Best Local Similarity   48.7%;  Pred. No. 0.18;
```

```
  Matches  115;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy   3157 ctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgcaa 3216
          ||  | || | || |||  || ||| ||||        || || | |||     | |
Db     70 ctggctgagagacccggagagccaggagccgcgggcggggaggcagaagggccggagggg 129

Qy   3217 agccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagacagaaagcgttgcc 3276
          |||   | |||   |   ||| |  || |||  ||||| |  |||   | | | |  | |
Db    130 agcgagggcgcagaggaggcgccgaggggcgccgccgctgtgaaggaggcaggaggcggc 189

Qy   3277 gaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcggaggaagagaaa 3336
          |  || | | ||    | |  |  ||       ||||  | |  |||| | |  ||||
Db    190 gggccagacaggggcccggaggccgaggcgcggggcacgaggggggcgcacggcgagact 249

Qy   3337 aaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaagcggaaac 3392
           |    | | |||  |     ||  | ||   || |   ||  | |  | ||| ||
Db    250 gaggccgaggagggagccccggagggtgccgaggtgccccaaggaggggaggagac 305


RESULT  38
AAA81518
ID   AAA81518 standard; DNA; 35133 BP.
XX
AC   AAA81518;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_65 SEQ ID NO:65.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1396-1406; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 35133 BP; 8216 A; 9873 C; 8987 G; 8057 T; 0 other;


  Query Match             1.0%;  Score 43;  DB 21;  Length 35133;
  Best Local Similarity   46.2%;  Pred. No. 1.4;
  Matches  133;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy   3617 gcngcatccggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaa 3676
          |  |||  |    |||||  | | ||  | | |||     |   ||| |||      |||
Db   2985 gacgcaaacccggcaccagccgccacgttacccaacgccgcgaagccgaccaagtcgaaa 3044

Qy   3677 ccgacctgcgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgt 3736
          |  |    ||  | |||      | |  | |  |     |     | ||   ||||
Db   3045 tcctctccggcgtattcgaaggcaaaaccaccggcacgcccatcgccctcttaatccgca 3104

Qy   3737 tttcgcacaaccggaccgaaaacancttcgacgacggcatcggcaactcggcacggcttg 3796
           | |  || | || | | || ||  |  | ||  || || | ||  |    |  | |  |
Db   3105 ataccgaccagcgcagcaaagactacggcaacatcgccaccagcttccgccccggccacg 3164

Qy   3797 cccacggcgccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgg 3856
          || ||    ||  ||      |||||||||   ||   | |  |  ||||  ||| |
Db   3165 ccgactatacctattggcacaaatacggcacgcgcgactaccgcggcggcggcaggagct 3224

Qy   3857 gcgcgggttttagcagcggcantctntcagacggcatcggaggcaaaa 3904
          |||  ||   | |  || |    |  | | |||   ||  | |||||
Db   3225 ccgcccgtgaaaccgccgcccgcgttgctgccggagccgttgccaaaa 3272


RESULT  39
AAF21611
ID   AAF21611 standard; DNA; 349980 BP.
XX
AC   AAF21611;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US05928.
XX
PR   30-APR-1999;   99US-0132068.
PR   08-OCT-1999;   99WO-US23573.
PR   28-FEB-2000; 2000GB-0004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M,  Hickey E,  Peterson J,  Tettelin H,  Venter JC,  Masignani V;
PI   Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V,  Rappuoli R;
PI   Frazer CM,  Grandi G;
XX
DR   WPI; 2000-647603/62.
```

```
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections -
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of
CC   Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC   to AAF21613 represent fragments of the NMB genomic sequence, as the
CC   sequence was too long to go in a record on its own it was split into 8
CC   sequences which overlap each other at the beginning and end of each
CC   sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC   the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC   the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC   Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC   AAF21606 represent PCR primers which are used in the exemplification of
CC   the present invention. The NMB genome and fragments from it have
CC   antibacterial activity, and can be used in vaccines and gene therapy.
CC   Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC   proteins can be used in compositions for treating or preventing infection
CC   due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC   presence of Neisserial bacteria or of antibodies raised to Neisserial
CC   bacteria. Computers, computer memory, computer storage medium or computer
CC   databases can be used in a search to identify open reading frames (ORFs)
CC   or coding sequences within the NMB genome. The DNA sequences provide
CC   further opportunities to find antigenic or immunogenic proteins which are
CC   more effective in vaccines than the outer membrane proteins currently
CC   used.
XX
SQ   Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 other;


  Query Match             1.0%;  Score 43;  DB 21;  Length 349980;
  Best Local Similarity   46.2%;  Pred. No. 5.3;
  Matches  133;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy    3617 gcngcatccggnacaccaaacactaccgttcgcaagatttccgcgcctaccgccaacaaa 3676
           |  |||  |    |||||  | | ||  | | |||      |   ||| |||      |||
Db  256301 gacgcaaacccggcaccagccgccacgttacccaacgccgcgaagccgaccaagtcgaaa 256360

Qy    3677 ccgacctgcgccaaatcggtatgcagaaaaacctcggcagcgggcgcgtcggcatcctgt 3736
             |  |    ||  | |||      | |  | |  |     |     | ||   ||||
Db  256361 tcctctccggcgtattcgaaggcaaaaccaccggcacgcccatcgccctcttaatccgca 256420

Qy    3737 tttcgcacaaccggaccgaaaacanctcgacgacggcatcggcaactcggcacggcttg 3796
           | |  || | || | | || ||  |  | ||  || || | ||  |    |  | |  |
Db  256421 ataccgaccagcgcagcaaagactacggcaacatcgccaccagcttccgccccggccacg 256480

Qy    3797 cccacggcgccgttttcgggcaatacggcatcggcaggttcgacatcggcatcagcacgg 3856
           || ||    ||  ||     |||||||||   ||   | |  |  ||||  ||| |
Db  256481 ccgactatacctattggcacaaatacggcacgcgcgactaccgcggcggcggcaggagct 256540

Qy    3857 gcgcgggttttagcagcggcantctntcagacggcatcggaggcaaaa 3904
           |||  ||   | |  || |    |  | | |||   ||  | |||||
Db  256541 ccgcccgtgaaaccgccgcccgcgttgctgccggagccgttgccaaaa 256588


RESULT  40
AAA81489
ID   AAA81489 standard; DNA; 837096 BP.
XX
AC   AAA81489;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US23573.
XX
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM,  Hickey E,  Peterson J,  Tettelin H,  Venter JC;
PI   Masignani V,  Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 629-865; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance:
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 837096 BP; 207534 A; 227065 C; 205215 G; 197280 T; 2 other;


  Query Match             1.0%;  Score 43;  DB 21;  Length 837096;
  Best Local Similarity   60.0%;  Pred. No. 8.9;
  Matches  311;  Conservative    0;  Mismatches  204;  Indels    3;  Gaps    1;

Qy    1171 tacggcaacggcaaactcatcttatcaaacaacatcaaccaaggcgcgggcggtttgtat 1230
           | || ||||   | ||| ||  |  || |||| ||||||||||||||| ||||  ||| |
Db  672949 ttcgacaacaaaacactgatgctggcagacaatatcaaccaaggcgcaggcgcattgcag 673008

Qy    1231 tttgaaggtgattttacggtctcgcctgaaaacaacgaaacgtggcaaggcgcgggcgtt 1290
           || ||  |  | || || |||     | ||||| ||   || |||||||| || ||||||
Db  673009 ttcgacagcaacttcaccgtcgtcggtaaaaaccac---acatggcaaggtgcaggcgtt 673065

Qy    1291 catatcagtgaagacagtaccgttacttggaaagtaaacggcgtggcaaacgaccgcctg 1350
             |    || | ||    |||    ||| |||| | |  |     |  |||||| ||
Db  673066 atcgtagccgacggcaaacgcgtcttctggcaagtcagcaaccccaaaggcgaccggctc 673125

Qy    1351 tccaaaatcggcaaaggcacgctgcacgttcaagccaaaggggaaaaccaaggctcgatc 1410
           |||||| | |||  |||||||||   ||   | |   ||||    ||||| |||   |||
Db  673126 tccaaactgggcgcaggcacgcttatcgccaacggacaaggcatcaaccagggcgacatc 673185

Qy    1411 agcgtgggcgacggtacagtcattttggatcagcaggcagacgataaaggcaaaaaacaa 1470
```

```
          ||| | || || || || ||| |  | |  ||  | || |      | ||||  ||||||
Db 673186 agcatcggggaaggcactgtcgtactcgcccaaaaagctgcttcagacggcagcaaacaa 673245

Qy   1471 gcctttagtgaaatcggcttgntcagcggcaggggtacggtgcaactgaatgccgataat 1530
          || || |   || ||||| |   |||||||||||| ||||     ||    | |    |
Db 673246 gcattcaaccaagtcggcatcaccagcggcaggggcacggccgtcctcgccgacagccag 673305

Qy   1531 cagttcaaccccgacaaactctatttcggctttcgcggcggacgtttggatttaaacggg 1590
          ||  |||| ||||| || ||||||||||||||  | ||||||||  | ||  | |||||
Db 673306 caaatcaaacccgaaaacctctatttcggcttcaggggcggacggctcgacctcaacggc 673365

Qy   1591 cattcgctttcgttccaccgtattcaaaataccgatgaaggggcgatgattgncnatcat 1650
           |    ||| | ||   || ||| |   || | || |  || |||   || | | ||||
Db 673366 aacaaccttgccttttacccatatccgccatgcggacggcggcgcgcaaatcgtcaatcac 673425

Qy   1651 aatgccacaacaacatccaccgttaccattacagggaa 1688
          ||  |      | |  | ||  | ||  | || || ||
Db 673426 aaccctgaccaagccgcgacactgacgctgaccggcaa 673463


RESULT  41
AAA47543/c
ID   AAA47543 standard; DNA; 1089 BP.
XX
AC   AAA47543;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Primate GPR2 sequence.
XX
KW   Cutaneous T-cell attracting chemokine; CTACK; skin; cell movement;
KW   migration; vasoactive intestinal contractor; Vic; GPR2; agonist;
KW   antagonist; antibody; immunological condition; mutein; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1089
FT                   /*tag= a
FT                   /product= GPR2
XX
PN   WO200038713-A1.
XX
PD   06-JUL-2000.
XX
PF   23-DEC-1999;   99WO-US30819.
XX
PR   24-DEC-1998;   98US-0113858.
PR   27-MAY-1999;   99US-0322580.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Wang W,  Oldham ER,  Soto H,  Lui Y,  Hudak SA,  Homey B,  Morales JM;
PI   Kellermann S,  McEvoy LM,  Zlotnik A;
XX
DR   WPI; 2000-465633/40.
DR   P-PSDB; AAB01448.
XX
PT   Modulating cell movement within the skin, useful for treating
PT   immunological skin conditions or diseases comprises administering T
PT   cell-attracting chemokine or vasoactive intestinal contractor chemokine
PT   agonists or antagonists
XX
PS   Example 3; Page 62-64; 79pp; English.
XX
CC   Modulating movement of a cell within or to the skin of a mammal can
CC   be achieved by administering an antagonist or agonist of cutaneous T
CC   cell-attracting chemokine (CTACK) or vasoactive intestinal contractor
CC   (Vic) chemokine. The antagonist is selected from a mutein of natural
CC   CTACK or Vic, an antibody which neutralises CTACK or Vic or an
CC   antibody which block GPR2 ligand binding. The CTACK or Vic agonists
CC   or antagonists are useful for treating medical conditions or diseases
CC   associated with immunological conditions of the skin.
XX
SQ   Sequence 1089 BP; 127 A; 409 C; 347 G; 206 T; 0 other;


  Query Match             1.0%;  Score 42.6;  DB 21;  Length 1089;
  Best Local Similarity   44.9%;  Pred. No. 0.23;
  Matches  153;  Conservative    0;  Mismatches  188;  Indels    0;  Gaps    0;

Qy   3155 agctttccgacaaactcggcaaggcagaagccaaaaaacaggcggaaaaagacaacgcgc 3214
          ||| | |  |    || |   | ||  | ||| |  ||    |  || || |  |  | ||
Db    971 AGCCTCCGCAGGTCCTGGCGGAAGCGCAGGCCCAGGAAGGCGTAGAGAACGGGATTGAGG 912

Qy   3215 aaagccttgacgcgctgattgcggccgggcgcgatgccgccgaaaagacagaaagcgttg 3274
            |     | || |      || || || | | |       ||| |     |     | ||
Db    911 CCACAGCGGGCGAGGGCCAAGCCGCTGGTCACCAGCAGTGCGACATCCTTGCGTTTGCTG 852

Qy   3275 ccgaaccggcccggcnggcaggcggggaaaatgtcggcattatgcaggcggaggaagaga 3334
            |    | || | ||  ||  || | |  |  |||||| ||| |||  | |||  |||
Db    851 GCAGGGCAGCTCCGCTCGCGCGCAGCCAGTAGATCGGCAGTATCCAGCAGCAGGGCGAGG 792

Qy   3335 aaaaacgggtgcaggcggataaagacagcgcnttggcgaaacagcgcgaagcggaaaccc 3394
            | ||  || |  | |  | ||  ||        ||  || | ||| ||| |   |
Db    791 CTGTAGGGCAGCTGCAGCACCACGAAGGCCGCCACCAGAGCCACCACGACGCGCAGCGCA 732

Qy   3395 ggccggntaccaccgccttcccccgcgcccgcngcgcccgccgggatttgccgcaaccgc 3454
            |||||    |    ||     | ||   |   | ||   || | |              |
Db    731 CGCCGGCGCTCGGGCCCCCTGGCGGCCAGCAGCGTGCGGCCCAGAAGCGCGTAGCAGGCT 672

Qy   3455 agccccaaccgcaacctcaaccccaaccgcagcgcgacctg 3495
          | |   |  | || |  || | | || | ||| || |||||
Db    671 ACCATGACGCCCAGCGGCAGCGCGAAGCCCAGGGCCACCTG 631


RESULT  42
AAV64545
ID   AAV64545 standard; DNA; 400 BP.
XX
AC   AAV64545;
XX
DT   27-JAN-1999  (first entry)
XX
DE   M. tuberculosis immunogenic polypeptide XP22 5'-end DNA.
XX
KW   Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW   vaccine; pharmaceutical; infection; diagnosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9816646-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US18293.
XX
PR   13-MAR-1997;   97US-0818112.
PR   11-OCT-1996;   96US-0730510.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Houghton R,  Lodes MJ;
PI   Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1998-261042/23.
XX
PT   Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT   to develop products for the detection of M. tuberculosis infection
PT   and for diagnosis, treatment and prevention of tuberculosis
XX
PS   Claim 4; Page 175; 230pp; English.
XX
CC   This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC   tuberculosis (MT) antigen which can be used in a method for inducing
CC   protective immunity against tuberculosis (TB). This sequence can be
```

```
CC   formulated into vaccines and/or pharmaceutical compositions for
CC   immunising against M. tuberculosis infection or may be used for the
CC   diagnosis of tuberculosis.
XX
SQ   Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;


  Query Match             1.0%;  Score 42.2;  DB 19;  Length 400;
  Best Local Similarity   45.8%;  Pred. No. 0.17;
  Matches  140;  Conservative    0;  Mismatches  166;  Indels    0;  Gaps    0;

Qy    3753 cgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtttt 3812
           ||  |||  |  || || ||||| | ||  | |   | |   | |  |  |||||
Db      92 cggcaacgccggcgccggcggcaccagcttcacccaaggcgccgacggcaacgccggcaa 151

Qy    3813 cgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcag 3872
           |||     |||||   | | |   | ||  ||| | | ||     ||  |     | ||
Db     152 cggcggtgacggcggggtcggcggcaacggcggaaacggcggaaacggcgcagacaacac 211

Qy    3873 cggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcat 3932
           |  ||      | | ||||| |  ||||     | || |   ||    |     ||| |
Db     212 caccaccgccgccgccggcaccacaggcggcgacggcggggccggcggggccggcggaac 271

Qy    3933 tcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaac 3992
             | |     | | ||||    |  |||   ||||  | |||    || |||||  |
Db     272 cggcggaaccggcggagccgccggcaccggcaccggcggccaacaaggcaacggcggcaa 331

Qy    3993 gcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccgg 4052
             ||     || |   ||||  |    | ||  |||    |        | || | ||||
Db     332 cggcggcaccggcggcaaaggcggcaccggcggcgacggtgcactctcaggcagcaccgg 391

Qy    4053 tcttgc 4058
           |  |||
Db     392 tggtgc 397


RESULT  43
AAV44436
ID   AAV44436 standard; DNA; 400 BP.
XX
AC   AAV44436;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen XP22 5' DNA.
XX
KW   Tuberculosis; infection; diagnosis; antigen; XP22; ss.
XX
OS   Mycobacterium tuberculosis strain Erdman.
XX
PN   WO9816645-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US18214.
XX
PR   13-MAR-1997;   97US-0818111.
PR   11-OCT-1996;   96US-0729622.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Campos-Neto A,  Dillon DC,  Houghton R,  Lodes MJ;
PI   Reed SG,  Skeiky YAW,  Twardzik DR,  Vedvick TS;
XX
DR   WPI; 1998-251292/22.
XX
PT   New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT   to develop products for the detection of M. tuberculosis infection
PT   and diagnosis of tuberculosis
XX
PS   Claim 4; Page 185-186; 250pp; English.
XX
CC   This is the 5' region of DNA coding for an antigenic portion of
CC   Mycobacterium tuberculosis antigen XP22; 3' DNA is provided in
CC   AAV44437.  XP22 DNA was isolated from a M. tuberculosis strain Erdman
CC   genomic DNA expression library using sera from patients having
CC   extrapulmonary tuberculosis.  It bears no similarity to known
CC   sequences.  The invention relates to methods and compositions for
CC   diagnosing tuberculosis.  It provides polypeptides (see
CC   AAW64291-W64379) comprising an antigenic portion of a soluble M.
CC   tuberculosis antigen, or an immunogenic portion of a M.
CC   tuberculosis antigen, as well as DNA sequences encoding such
CC   polypeptides, recombinant expression vectors and transformed or
CC   transfected host cells.  Also claimed are methods and diagnostic
CC   kits for detecting M. tuberculosis infection in a patient using
CC   these polypeptides, antibodies or oligonucleotide probes and
CC   primers.
XX
SQ   Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;


  Query Match             1.0%;  Score 42.2;  DB 19;  Length 400;
  Best Local Similarity   45.8%;  Pred. No. 0.17;
  Matches  140;  Conservative    0;  Mismatches  166;  Indels    0;  Gaps    0;

Qy    3753 cgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtttt 3812
           ||  |||  |  || || ||||| | ||  | |   | |   | |  |  |||||
Db      92 cggcaacgccggcgccggcggcaccagcttcacccaaggcgccgacggcaacgccggcaa 151

Qy    3813 cgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcag 3872
           |||     |||||   | | |   | ||  ||| | | ||     ||  |     | ||
Db     152 cggcggtgacggcggggtcggcggcaacggcggaaacggcggaaacggcgcagacaacac 211

Qy    3873 cggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcat 3932
           |  ||      | | ||||| |  ||||     | || |   ||    |     ||| |
Db     212 caccaccgccgccgccggcaccacaggcggcgacggcggggccggcggggccggcggaac 271

Qy    3933 tcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaac 3992
             | |     | | ||||    |  |||   ||||  | |||    || |||||  |
Db     272 cggcggaaccggcggagccgccggcaccggcaccggcggccaacaaggcaacggcggcaa 331

Qy    3993 gcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccgg 4052
             ||     || |   ||||  |    | ||  |||    |        | || | ||||
Db     332 cggcggcaccggcggcaaaggcggcaccggcggcgacggtgcactctcaggcagcaccgg 391

Qy    4053 tcttgc 4058
           |  |||
Db     392 tggtgc 397


RESULT  44
AAZ19346
ID   AAZ19346 standard; DNA; 400 BP.
XX
AC   AAZ19346;
XX
DT   05-NOV-1999  (first entry)
XX
DE   M. tuberculosis antigen 5' XP22 DNA sequence.
XX
KW   Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW   immunotherapy; diagnosis; immunisation; vaccine; infection;
KW   immune response; skin test; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9942076-A2.
XX
PD   26-AUG-1999.
XX
PF   17-FEB-1999;   99WO-US03268.
XX
PR   05-MAY-1998;   98US-0072967.
PR   18-FEB-1998;   98US-0025197.
XX
PA   (CORI-) CORIXA CORP.
```

XX
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1999-527409/44.
XX
PT New antigens from Mycobacterium tuberculosis useful in diagnostic
PT skin tests and protective or therapeutic vaccines or compositions
XX
PS Claim 4; Page 170; 299pp; English.
XX
CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC are vaccines and fusion protein containing M. tuberculosis Ag's.
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC other polypeptides fragments, can be used in pharmaceutical compositions
CC or vaccines to generate a protective or therapeutic immune response to
CC M. tuberculosis and as reagents in skin tests for diagnosis of
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC by, T, B or natural killer cells and/or macrophages in
CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
CC AAY39225 are used in the exemplification of the present invention.
XX
SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;

Query Match 1.0%; Score 42.2; DB 20; Length 400;
Best Local Similarity 45.8%; Pred. No. 0.17;
Matches 140; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

```
Qy   3753 cgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtttt 3812
          ||  |||  |  || || ||||| | ||  | |   | |    | |  |  |||||
Db     92 cggcaacgccggcgccggcggcaccagcttcacccaaggcgccgacggcaacgccggcaa 151

Qy   3813 cgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcag 3872
          |||     |||||   | | |   | ||  ||| | | ||     ||  |     | ||
Db    152 cggcggtgacggcggggtcggcggcaacggcggaaacggcgcggaaacggcgcagacaacac 211

Qy   3873 cggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcat 3932
          |  ||      | | ||||| |  ||||     | || |   ||    |     ||| |
Db    212 caccaccgccgccgccggcaccacaggcggcgacggcggggccggcggggccggcggaac 271

Qy   3933 tcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaac 3992
            | |     | | ||||    |  |||   ||||  | |||    || |||||  |
Db    272 cggcggaaccggcggagccgccggcaccggcaccggcggccaacaaggcaacggcggcaa 331

Qy   3993 gcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccgg 4052
            ||     || |   ||||  |    | ||  |||    |         | || | ||||
Db    332 cggcggcaccggcggcaaaggcggcaccggcggcgacggtgcactctcaggcagcaccgg 391

Qy   4053 tcttgc 4058
          |  |||
Db    392 tggtgc 397
```

RESULT 45
AAZ19134
ID AAZ19134 standard; DNA; 400 BP.
XX
AC AAZ19134;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis recombinant antigen DNA encoding 5' XP22.
XX
KW Antigen; diagnosis; detection; infection; antibody; immunisation;
KW vaccine; immunity; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942118-A2.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US03265.
XX
PR 05-MAY-1998; 98US-0072596.
PR 18-FEB-1998; 98US-0024753.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1999-527416/44.
XX
PT New polypeptide comprising antigenic portions of M. tuberculosis
XX
PS Claim 4; Page 215-216; 323pp; English.
XX
CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.
XX
SQ Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;

Query Match 1.0%; Score 42.2; DB 20; Length 400;
Best Local Similarity 45.8%; Pred. No. 0.17;
Matches 140; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

```
Qy   3753 cgaaaacancttcgacgacggcatcggcaactcggcacggcttgcccacggcgccgtttt 3812
          ||  |||  |  || || ||||| | ||  | |   | |    | |  |  |||||
Db     92 cggcaacgccggcgccggcggcaccagcttcacccaaggcgccgacggcaacgccggcaa 151

Qy   3813 cgggcaatacggcatcggcaggttcgacatcggcatcagcacgggcgcgggttttagcag 3872
          |||     |||||   | | |   | ||  ||| | | ||     ||  |     | ||
Db    152 cggcggtgacggcggggtcggcggcaacggcggaaacggcggaaacggcgcagacaacac 211

Qy   3873 cggcantctntcagacggcatcggaggcaaaatccgccgccgcgtgctgcattacggcat 3932
          |  ||      | | ||||| |  ||||     | || |   ||    |     ||| |
Db    212 caccaccgccgccgccggcaccacaggcggcgacggcggggccggcggggccggcggaac 271

Qy   3933 tcaggcacgataccgcgccggtttcggcggattcggcatcgaaccgtacatcggcgcaac 3992
            | |     | | ||||    |  |||   ||||  | |||    || |||||  |
Db    272 cggcggaaccggcggagccgccggcaccggcaccggcggccaacaaggcaacggcggcaa 331

Qy   3993 gcgctatttcgtccaaaaagcggattaccgctacgaaaacgtcaatatcgccacccccgg 4052
            ||     || |   ||||  |    | ||  |||    |         | || | ||||
Db    332 cggcggcaccggcggcaaaggcggcaccggcggcgacggtgcactctcaggcagcaccgg 391

Qy   4053 tcttgc 4058
          |  |||
Db    392 tggtgc 397
```

Search completed: July 1, 2002, 08:01:13
Job time: 133106 sec